Current Awareness Bulletin

of

SCHOLARLY ARTICLES PUBLISHED BY

Faculty, Students and Alumni

~ November 2012 ~

DELHI TECHNOLOGICAL UNIVERSITY CENTRAL LIBRARY
(formerly Delhi College of Engineering, Bawana Road, DELHI)

PREFACE

This is the first Current Awareness Bulletin Service started by Delhi Technological University Library. The aim of the bulletin is to compile, preserve and disseminate information published by the Faculty, Students and Alumni for mutual benefits. The bulletin also aims to propagate the intellectual contribution of DTU as a whole to the academia. It contains information resources available in the internet in the form of articles, reports, presentation published in international journals, websites, etc. by the faculty and students of Delhi Technological University in the field of science and technology. The publication of Faculty and Students, which are not covered in this bulletin, may be because of the reason that the full text either was not accessible or could not be searched by the search engine used by the library for this purpose. To make the bulletin more comprehensive, the learned faculty and Students may provide their uncovered publication to the library either through email or in CD, etc.

This issue contains the information published during November 2012. The arrangement of the contents is alphabetical wise starting from A-Z. The Full text of the article, which is either subscribed by the University or available in the web, is provided in this Bulletin.

CONTENTS

14. Regressive Structures for Computation of DST-II and Its Inverse by ***@Priyanka Jain*** and Anamika Jain.

15. Software Maintainability Prediction using Machine Learning Algorithms by ****Ruchika Malhotra*** and Anuradha Chug.

16. Thermodynamic Analysis for improbvement in thermal performance of a simple gas turbine cycle through retrofitting techniques(inlet air evaporativecooling steam injection and combined IAC and STIG) by ****Shyam Agarwal and *R.S. Mishra.***

17. Yagi-Uda Antenna Design Optimization for Maximum Gain using different BBO Migration Variants by ***#Satvir Singh*** *and Gagan Sachdeva.*

***** **Faculty**
@ **Students/Research Scholars**
**Alumni**

Proceedings of the National Conference on
Trends and Advances in Mechanical Engineering,
YMCA University of Science & Technology, Faridabad, Haryana, Oct 19-20, 2012

A GENERIC MODEL OF MULTI-ECHELON REVERSE LOGISTICS NETWORK FOR PRODUCT RETURNS

S. Bansal[1], A.Jayant[1], P. Gupta[1], S. K. Garg[2]
[1]Department of Mechanical Engineering, Sant Longowal Institute of Engineering and Technology, Longowal, Sangrur – 148106
[2]Department of Mechanical Engineering, Delhi Technological University, Delhi-110042

Abstract
Rapid technology advances in turbulent Indian business environment have shortened the lifecycle of white goods, resulting in the increasing number of discarded products in recent years. Due to the growing environmental concerns, several state governments have passed new regulations in order to reduce the amount of waste stream generated by mass consumption of the products in the society, to divert the discarded/End-of – Life (EOL) products from landfills, and to dispose the retired electronic & mechanical assembly based products properly. As a result, an effective reverse logistics infrastructure is required to support the product recovery activities. In this research, a noble approach for designing reverse logistics infrastructure by private-government partnership model is presented. Finally, discussion, recommendation and insight information in operating reverse logistics real business environment is analyzed and provided.

1. Introduction
Supply chain management (SCM) can be considered as a key component of competitive strategy to enhance organizational productivity, performance and profitability [7]. In the recent past, there has been a surge in research that examined the impact of supply chain integration on firm performance. Most SCM publications concern mainly procurement production, extending the concept beyond the point of sale is rare. Recently, increased need has been recognized to extend SCM issues beyond the point of sale in industrial manufacturing. Hence, research field of managing supply chains has been enlarged by tasks referring to the product utilization phase (e.g. service, maintenance, and others) and to the end-of-life phase (e.g. product recovery, refurbishing or recycling). Conceptually speaking, these additional tasks have been complementary traditional supply chains to closed-loop supply chains [12].

Sustainability initiatives brought increasingly growing number of countries across EU and Eastern Asia to enact legislations that would demand manufacturers to assume higher responsibilities on their end-of-life products [35]. Sustainability is becoming one of the most desired and highly prized goals of modern industrial operations and environmental management as the deterioration of natural environment becomes increasingly more concerned. International Union for the Conservation of Nature and Natural Resources, the Global Tomorrow Coalition, and the World Resources Institute establish sustainability as a desired goal of environmental management, development and international cooperation. The term, "sustainability," issued in numerous disciplines and is defined in many ways according to the context to which it is applied and whether its use is based on an ecological, social, or economical perspective. IUCN defines sustainability as improving the quality of human life while living within the carrying capacity of supporting eco-systems. Although conceptualization of sustainability may differ among different interest groups, the World Commission on Environment and Development defines sustainable development, as 'development that meets the needs of the present without compromising the ability of the future generations to meet their own needs [6]. In many Western European countries, "Green" parties have been initiated to deliver environmental concerns due to industrial and operational wastes into public, social and political action. In response to globally growing concerns for sustainability, many durable product manufacturers began to launch programs that would both reduce operational wastes and advocate environmental safety. The intent of the 'product take-back' laws is to pressurize durable product manufacturers to pursue sustainable development and to transform it into business practices that would promote environmental welfare, while avoiding increasingly growing waste management cost charged by municipal governments. In addition, higher customer expectations on manufacturers' environmental responsibility have also compelled manufacturers to assume increased responsibility with regards to placing their products on the market. 'Product take-back' targets a wide variety of manufacturers of batteries, automobiles, waste packaging, and electrical or electronic products. Instead of filling landfills, more manufacturers are urged to take back their products for reassembling, repackaging, remanufacturing, or component recycling before redistributing to the market. Value recovery process of returned products consists of several sequential activities: collection, evaluation, disassembly, capture of recyclable components, and disposal of residuals as hazardous wastes [11].

2. Need of economically and environmentaly viable rl frame work

Despite growing participation within industries, most value recovery processes still remain small, independent and highly fragmented [34]. To strategize cost efficient product take-back plan, there has been growing interest in the development of reverse logistics that drives reverse flow of returned products from the end customers back to the original equipment manufacturers. Efficient planning and execution of reverse logistics would provide firms a competitive edge in the development of sustainable, yet profit-generating, business strategies. Sound strategy and execution of reverse logistics would promote not only economic, but also environmental benefits as value of returned products should be counted towards savings of raw material and labor. While reverse logistics do not promise guaranteed savings, many have reported noticeable benefits: 40% less overall cost, 33% less inventory usage, and 44% higher customer satisfaction .From environmental viewpoint, reverse logistics make significant contribution towards reduction of hazardous waste, alleviation of landfill saturation and preservation of scarce raw materials [12]. Reverse logistics take fundamentally different approach from forward logistics having characteristics of highly fragmented return quantities, multiple return channels, complex transportation routing, higher level of expected serviceability for multiple Clients and variety of disposition options. Due to such characteristics, realization or execution of reverse logistics often entail many new challenges. Two major challenges of reverse logistics network design would include cost of value recovery process and low return rates from customers. Recent research reported the cost of reverse logistics accounts for nearly 44% of entire product take-back process [41]. Additionally, Green peace's survey in 2007 revealed that many manufacturers struggle to achieve beyond 20 percent of product return rate. Challenges in product take-back processes entail careful evaluation of aforementioned two key issues of reverse logistics network design in order to minimize the total operating cost, while promoting higher customer product return frequency.

However, the reverse flow of products from consumers to upstream business has not received much interest [13]. Yet, reverse logistics is a big business opportunity. According to the survey in 1999 that reverse logistics executive council the cost of handling, transporting and determining the disposition of returned products was $35 billion annually for U.S firms [14]. In 2000 remanufacturing in the U.S. was a #35 billion per year industry [15]. Up to now, rates are still increasing. In India e-waste generation from electronics and computers industries is approximate 1050 tons and if we count imported used products from developed countries then this figures goes up to three times. From abroad used computers and electronic goods send to India by illegal practices and received by big scrap dealers. The scrap dealers take outs useful components from used products like circuit boards, switches, condenser, capacitor, batteries, transformers, copper wires, aluminums wires and other precious metals by unscientific techniques and unsafe methods. By this process various dangerous gases and metal particle like cadmium, mercury, bromine flame, poly- chlorination, Bi-finials are mixed in the environment and causes for cancer, respiratory and brain related diseases in the society.

3. SAFE HANDLING OF E-WASTE

Today world is facing problem of e-waste. The life of this e-waste is very long and it is not biodegradable and remains in the environment for long period, Product life cycle has been very short and day to day companies are launching new advance products due to developments of advance technologies in the world. That is main reason for the creation of e-waste in the society. In India more than 10 million computers are under use and more than 2 million computers are outdated. At present in our country here is no rule or directives from government regarding retreatment and recycling of e-waste. To save the environment and society there is need to develop an economically viable and safe practice of reverse logistics/recycling model. This model may be helpful to rescue of products/components in environment friendly manner. We hereby developed a conceptual holistic generic frame work of forward and reverse supply chain networks as shown in figure 1.

3.1 Secondary Markets

To generally conceptualize, reverse logistics is the process of retrieving the product from the end consumer for the purposes of capturing value or paper disposal. Activities include transportation, warehousing, distribution and inventory management. Transportation is usually the largest component of reverse logistics costs. Reverse logistics services include product returns, source reduction , recycling, materials substitution, reuse of materials, waste disposal, refurbishing, repair and remanufacturing [18] Reverse logistics -and reverse logistics research- has traditionally emphasized green logistics i.e. the use of environmentally conscious logistics strategies [18,19]. While environment aspects of reverse logistics are critically important, many firms are also recognizing the economic impact of reverse logistics [20] Practically all business must deal with returns of some natures because of issues such as marketing returns (i.e. customers change their minds or find the product unacceptable), damage or quality problems, overstocks, or, merchandise that is brought back for repairs, refurbishing, or remanufacturing. NOREK (2002) provides an indication of the sheer volume of returns generated in many

companies. He notes that returns range from 3% to as high as 50% of total shipments across all industries; various industry studies put the true costs of returns at 3-5% of sales [21].

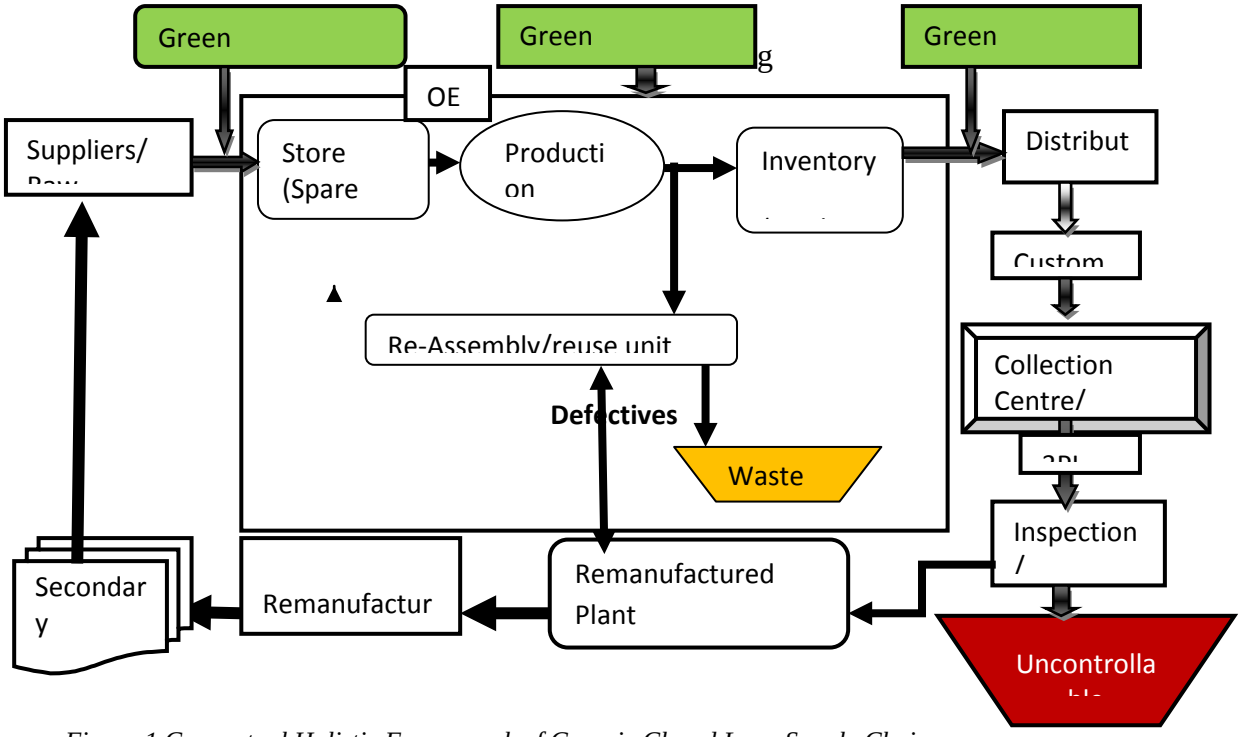


Figure 1 Conceptual Holistic Frame work of Generic Closed Loop Supply Chain

In brief, the management of the reverse flows is an extension of traditional supply chains with used product or material either returning to reprocessing organizations or being discarded. Reverse supply chain management is defined as the effective and efficient management of the series of activities required to retrieve a product from a customer and either dispose of it or recover value. The importance of studying reverse supply chains has increased in recent years for several reasons:

Sales opportunities in secondary and global markets have increased revenue generation from previously discarded products.

1) End-of-life take-back laws have proliferated over the past decade both in the European Union and in the United States, requiring businesses to effectively manage the entire life of the product [22, 23].
2) Consumers have successfully pressured business to take responsibility for the disposal of their products that contain hazardous waste [24].
3) Landfill capacity has become limited and expensive. Alternatives such as repacking, Re-manufacturing and recycling have became more prevalent and viable [25, 26]

In conclusion, become of effective reverse logistics in daily operations, firms can to foster a sustainable competitive advantage and increase revenues in a highly competitive market.

4. Green supply chain management operation

Some of the key challenges of GSCM such as integrating remanufacturing with internal operations, understanding the effects of competition among remanufacturers, integration product design, integrating remanufacturing and RL with supply chain design are discussed as under.

4.1 Green manufacturing and remanufacturing operation

This is a very important area within green operations. The techniques for minimum energy and resource consumption for flow systems in order to reduce the use of virgin materials are based on three fields of study: pinch analysis, industrial energy [10] and energy and lifestyle analysis. Logistics represent up to 95% of total costs (stock 1998 in recycling. Automobile, electronic, and paper recycling are the most common examples of product recovery the purpose of repair is to return used products to working order. The quality of repaired

products is generally lower than the quality of new products. The purpose of refurbishing is to bring used products up to a specified quality. Analysis of remanufacturing facilities for household appliances and automotive parts by [21] reveals that cleaning and repairing are the most critical steps in the re manufacturing process.

4.2 Organizational Size and Environmental Practice

It is seen whether resources and capabilities associated with different sized organizations play a role in adopting GSCM practices. Determining whether smaller organizations are adopting at greater, lesser one even equal rate as compared to medium and larger organizations for environmental practices sets the foundation for practical and research issues. D inferences among these organizations will influence different strategies that can be applied by supply chain and logistics partners, investors, professional organizations and government policy makers to aid smaller or larger organizations or both when seeking to adopt these GSCM practices and innovations. For example, if larger organizations are adopting practices earlier than their smaller counterparts. Then a diffusion mechanism through collaborative partnerships with smaller organizations may be a policy that should be encouraged.Yet, if all organizations seem to be lagging in a particular GSCM practice adoption, supply chains, umbrella professional groups or even regulatory agencies may play a larger role in diffusing these innovative practices.

4.3 Recent trends and examples in GSCM

In recent years, some organizations have begun relying on their supply chains to improve their business performance and create value for their end customers. Manufacturers also are calling on their suppliers more frequently to create innovative ideas that exploit new emerging technologies, and reduce costs during the design and development of their products [18]. In some instances, organizations are even relying on their suppliers to deliver state-of-the-art process technology that they cannot develop internally. Consequently, enterprises wishing to minimize their environmental impacts during product design are learning that their ability to do so often is dependent on their ability to manage their increasingly complex supplier relationships. For instance, in 2002, Hewlett-Packard established its Supply Chain Social and Environmental Responsibility Policy. The company also instituted a supplier code of conduct. Combined, these efforts have extended Hewlett-Packard's corporate social responsibility commitment by incorporating its global supply base and reducing its supply chain risks.

4.4 Capabilities for Adopting GSCM

There are numerous capabilities required to adopt GSCM, Organizations have to develop their knowledge-based competencies by guaranteeing the environmental quality of incoming goods.GSCM practices require organizations to have strong inventory control systems. These systems reduce redundant stock materials and unnecessary inputs in the production process. Organizations that rely on these systems should manage materials, productive capacity and other organizational information. At their core, GSCM rely on what on Deming's (1986) continuous improvement model.GSCM practices leverage continual improvement processes that reduce the impact of supplier inputs on the organization's final product. Collaboration across internal departments is essential to maintaining robust GSCM practices. For instance, in utilizing GSCM,an organization must coordinate its product design department with its marketing department and its suppliers in an effort to minimize waste and environmental impact at every node in the supply chain [18].However, traditional organizational structures generally are fragmented with purchasing departments operating separately from marketing and sales, and operations functioning independently from human resources, with each having their own goals.

4.5 Top Antecedents of GSCM

4.5.1Management Commitment

Implementation of GSCM practices in any manufacturing environment is a strategic decision, as it requires significant amount of time, effort and resources. Min et al. [21] mentioned that one of the major obstacles for implementing environmental policies is the lack of top management support. Kroon [11] proposed top-level support as one of critical elements for the successful implementation of GSCM.Zsidisin and Siferd (2001), Trowbridge (2001), and Rice (2003) mentioned that top management must be committed to complete environmental excellence.Hu and Hsu (2006) demonstrated analytically that top management support is the most important item for the successful implementation of GSCM practice in the Taiwanese electrical and electronics industries.

4.5.2 Government's Initiative

The government should ignite, encourage and promote the green activities carried out by the manufacturing supply chain.Zhu et al. (2005) mentioned that China has encouraged (pressured) GSCM practice adoption to help spur economic development. One can find literatures which also claim government regulation as the major driver of environment/green efforts of manufacturing companies (Green et al., 1996; Handfield et al., 1997; Walton et al., 1998; Eagan and Kaiser, 2002; Scupola, 2003; Lin, 2007; and Peng and Lin, 2008).The government manufacturing companies should become role models to others. They should come out with transparent legislation for environmental responsibility. Environmental regulations such as EuP (Eco-Design of Energy-Using Products), REACH (Registration, Evaluation, and Authorization of Chemicals), ELV (End of life Vehicle Directive), WEEE (Waste from Electrical and Electronic Equipment) and RoHS (Restrictions on Hazardous Substances) of the European Union are putting pressure on the companies for adopting green practices. One can find similar types of laws in other countries like China, Taiwan, Korea, Japan and the US. Such laws compel the manufacturing companies to closely look into the production processes and supplier selections. Within a few years, products will be sold in most parts of world under such legislation. The government should apply pressure so as to compel the companies to become green without any compromise.

4.5.3 External Pressures for Adopting GSCM

GSCM may be considered complimentary management practices relate to the institutional pressures that encourage their adoption**. Institutional pressures** persuade organizations to undertake similar strategic actions (Hoffman, 1997; Scott, 2001) to increase their external legitimization (DiMaggio and Powell, 1983; Hoffman and Ventresca, 2002). Regulatory pressures are often associated with an organization's decisions to adopt GSCM practices (Birett, 1998). These pressures arise from threats of non compliance penalties and fines (Davidson and Worrell, 2001) and requirements to publicly disclose information about toxic chemical releases(Konar and Cohen, 1997).For instance, regulatory changes in automotive paints have pressed car manufacturers to require their suppliers to reduce their use of regulated chemicals in the production process(Geffen and Rothenberg,2000). In addition to regulatory pressures**, Market Pressures** may influence an organization's decision to adopt on GSCM practices (Rao, 2002; Gupta and Piero, 2003). Over the last ten years, market actors have been placing greater pressures on organizations to consider their impacts on the natural environment (Hoffman, 2000). Overall, 15 percent of US consumers routinely pay more for green products, and another 15 percent seek green products if they do not cost more (Ginsberg and Bloom, 2004). While these findings suggest that markets are creating opportunities for environmental friendly organizations, the majority of consumers still are not influenced by a company's proactive environment practices. However, these same customers may be persuaded to change their purchasing decisions if a company violates environmental laws or emits high levels of toxins (Prakash, 2000). As a consequence, EMS and GSCM adoption may provide a vehicle for organizations to 'signal' to market participants that their environmental strategies adhere to or exceed generally accepted environmental standards. Doing so may lead to greater acceptance of the organization's strategic approach (DiMaggio and Powell, 1983) and insulate organizations from competitor's criticisms (King and Lenox, 2001). EMS and GSCM adoption also may help organizations develop an environmentally conscious reputation. Such a re-reputation may invite patronage from consumers and generate opportunities for business with other organizations that value these principles (Darnall and Carmin, 2005). Finally, organizations are subjected to **Pressures from the community** that includes environmental groups, community groups, media, labor unions and industry associations (Hoffman, 2000). Each of these groups can marshal public support for or against an organization's environmental performance (Clair, Milliman and Mitroff, 1995; Turcotte, 1995).

4.5.4 Green Procurement

Green procurement is defined as an environmental purchasing consisting of involvement in activities that include the reduction, reuse and recycling of materials in the process of purchasing. Besides green procurement is a solution for environmentally concerned and economically conservative business, and a concept of acquiring a selection of products and services that minimizes environmental impact Supplier selection: (1) purchase materials or parts only from "Green Partners" who satisfy green partner environment quality standards and pass an audit process in following regulations for the environment-related substances (2) consider suppliers who acquire ISO 14000, OHSAS18000 and/ or RoHS directives(3) select suppliers who control hazardous substances in company's standard lists and obtain green certificate achievements EPP is the act of purchasing products or services that have a less adverse effect on human health and the environment.

4.5.5 Green Manufacturing

Green manufacturing is defined as production processes which use inputs with relatively low environmental impacts, which are highly efficient, and which generate little or no waste or pollution. Green manufacturing can lead to lower raw material costs, production efficiency gains, reduced environmental and occupational safety

expenses, and improved corporate image. Activities in green manufacturing are: Hazardous substance control: (1) lead free-replace other substances such as bismuth, silver, tin, gold, copper (2) rinse parts with clean water instead of using chemicals and reuse water (3) quality control in inputs at vendor site and recheck before processing Energy-efficient technology: (1) reduce power consumption in products such as ramp load/unload technology in HDD (2) increase product lifespan resulting in higher efficiency and productivity.

4.5.6 Green Distribution

Green distribution consists of green packaging and green logistics. Packaging characteristics such as size, shape and materials have an impact on distribution because of their effect on the transport characteristics of the product. Better packaging along with rearranged loading patterns can reduce materials usage, increase space utilization in the warehouse and in the trailer, and reduce the amount of handling required. Activities in green distribution are Green packaging: (1) downsize packaging (2) use "green" packaging materials (3) cooperate with vendor to standardize packaging (4) minimize material uses and time to unpack (5) encourage and adopt returnable packaging methods (6) promote recycling and reuse programs. Green logistics/transportation: (1) deliver directly trouser site (2) use alternative fuel vehicles (3) distribute products together, rather than in smaller batches (4) change to modal shift.

4.5.7 Reverse Logistics

One of the collective solutions that industries have come up with is the development of the reverse logistics that focus on the value recovery of returned products for recycling or remanufacturing. Reverse logistics refers to the logistics management skills and activities involved in reducing, managing and disposing packages or products. Srivastava defines reverse logistics as "Integrating environmental thinking into supply chain management including product design, material sourcing and selection, manufacturing processes, delivery of the final product to the consumers as well as end-of-life management of the product after its useful life". A growing responsibility towards the environment and governmental regulations, and increasing awareness of valuable commercial opportunities in collecting, recycling, and reusing products and materials stimulate the development. One of the obvious challenges of reverse logistics is reverse distribution of goods and information; which fundamentally differs from that of forward logistics in terms of direction of material and information flow and their respective volume. Due to its difficulties in handling, reverse logistics cost exceeds $35 billion dollars per year for US companies. For above reasons, many companies treat reverse-logistics as a non-revenue-generating process which would often result in a very few resources allocated to this part of the supply chain. However, more and more firms now realize that reverse logistics is a business process by itself with growing attention towards sustainability and environmental responsibility. Hawken et al. envision economic benefits of as much as 90% through reduction of energy and materials consumption. Practice of reverse logistics entails a series of tasks to capture value of products returned for recycling.

Product acquisition to obtain the products from end-users

- Transshipment from point of acquisition to a point of disposition
- Testing, sorting, and disposition to determine products' economic attractiveness
- Refurbish to facilitate the most attractive economic options: reuse, repair, Remanufacture, recycle, or disposal
- Remarketing to create and exploit secondary markets

As reverse logistics fundamentally differ in many aspects of operations from forward logistics, strategic development of competitive reverse logistics entails careful evaluation, design, planning and control. Product acquisition would initiate at initial collection centers (ICPs) and consolidation would continue before reaching centralized return center (CRC) or manufacturer who would process remanufacturing.

5. Industry response towards reverse logistics

In many ways, industries have been focusing on maximizing financial or productive capital gain while consuming natural and social capital as needed. Global environmental awareness, however, have brought environment friendly or green initiatives in every aspect of product operations. Xerox's accomplishment of 'zero-waste to-landfill' engineering can be a very good example of 'cleaner production [15]. Increasingly many industries have adopted concepts of cleaner production and developed many strategic approaches and practices that increase re-manufacturability or recyclability of products or eliminate harmful wastes. Waste Electric and Electronic Equipment (WEEE) directive of the European Union; for instance, obliges manufacturers of electric and electronic equipment to assume extended responsibility by taking back equipments reached end-of-life state for re-processing and recovery.

Radical transformation did more than mere improvement of corporate images. The financial impact has been remarkable. 3M's 3P (also known as Pollution Prevention Pays) project has saved the company more than $1 billion in its first year by aggressively limiting harmful by-products and wastes. Kathy Reed of 3M noted "Anything not in a product in a product is considered a cost". Timberland's redesigned shoeboxes saved nearly 15% of virgin packaging material. AMD's modified 'wet processing' technology reduced the water usage from eighteen to less than now six gallons per minute. Besides many notable individual achievements, the sustainability issues must be dealt at supply chain managements' level as today's industries become more and more interdependent on one another in every aspect of product and service delivery. Efforts of environmental management and operations should no longer be limited to issues of localized product operations. Rather, it needs to be assessed in a higher level of Operations, which encompass production, transportation, consumption and post-disposal Disposition. Given such a significant and increasing level of attention toward issues related sustainable development, or sustainability, it is imperative to define sustainability on supply chain managements' level to discuss environmental as well as economic benefits as a whole.

6. Government and industry partnership

The problem of e-waste can be solved by joint efforts of industry and government. The government should ignite, encourage and promote the green activities carried out by the manufacturing supply chain. China has encouraged (pressured) GSCM practice adoption to help spur economic development. One can find literatures which also claim government regulation as the major driver of environment/green efforts of manufacturing companies. The government manufacturing companies should become role models to others. They should come out with transparent legislation for environmental responsibility. Environmental regulations such as EuP (Eco-Design of Energy Using Products), REACH (Registration, Evaluation, and Authorization of Chemicals), ELV (End of Life Vehicle Directive), WEEE (Waste from Electrical and Electronic Equipment) and RoHS (Restrictions on Hazardous Substances) of the European Union are putting pressure on the companies for adopting green practices. One can find similar types of laws in other countries like China, Taiwan, Korea, Japan and US. Such laws compel the manufacturing companies to closely look into the production processes and supplier selections. Within a few years, products will be sold in most parts of world under such legislation. The government should apply pressure so as to compel the companies to become green without any compromise. There is a need to make policies by the Indian government to handle the problem of huge e-waste generated by the producers/manufacturing industries. Municipalities can play lead role in the collection of waste/used/EOL products from the society with collaboration of industries and 3PL provider to optimize the use of EOL products. We hereby developed an integrated model of forward and reverse supply chain for efficient utilization of EOL products in India with the concept of government and private partnership strategy.

The structure of the presentation was based on functions that could be considered as drivers within the green supply chain/closed loop supply chain. These are Procurement, in-bound logistics, production, distribution and out-bound logistics, and reverse logistics procedure with joint collaboration of government and private partnership model.

7. Conclusion & Discussion

The underlying aim in considering the end-of-life phase of a product's life is to reduce impacts on the natural environment. The ultimate goal is sustainable development "meeting the needs of the present without compromising the ability of future generations to meet their own needs". The perspective of this work is on manufacturer involvement in managing end-of-life products with the infrastructure support of local government like, municipal Corporations, Nagar Councils, Village Councils and Third Party Logistics Service (3PL) provider. This study focuses on the Original Equipment Manufacturers (OEM) relationship with government agencies to handle the product returns

The generic model presented in the paper represents total perspective and look at the problem of used product return from an overall environmental or societal perspective. The strategic perspective to end-of-life management with government support has received very limited attention [12], especially the role of manufacturers which is expected to grow [14]. Waste collectors can be municipalities, third parties, or logistics service providers. Depending on their location in the world, end users may have to pay to dispose of the product. From waste collection the product will be sent to a landfill, an incinerator, or a recycling facility. Incineration means that energy is recovered from the product, whereas recycling refers to recovering material value from the product [8]. From the recycling facilities the recovered materials may end-up back in the original supply chain of the product or an alternative supply chain. Recycling facilities may in some cases be owned by manufacturers, as is frequently the case in Japan [10].

The government should apply pressure so as to compel the companies to become green without any compromise. There is a need to make policies by the Indian government to handle the problem of huge e-waste generated by the producers/manufacturing industries. As per present model theme the municipalities can play lead role in the collection of waste/used/EOL products from the society with collaboration of industries and 3PL provider to optimize the use of EOL products. We hereby developed an integrated model of forward and reverse supply chain for efficient utilization of EOL products in India with the concept of government and private partnership strategy.

References

1. Amato-McCoy, D.M., 2003. “Sears gets a return on returns”, Stores, Vol. 85 No. 7, pp.66.
2. Bidwell, R. and Verfaillie, H.A., 2000. Measuring Eco-efficiency: a Guide to Reporting Company Performance. Geneva: World Business Council for Sustainable Development.
3. Bradbury, H. and Clair, J.A., 1999. Promoting sustainable organizations with Sweden’s Natural Step. Academy of Management Executive, 13, 4, 63-74.
4. Chopra, S. Meindl, P. Supply Chain Management 3rd edition Harvard Business Press Chung, Chun-Jen, Wee, Hui-Ming, 2008. Green-component life-cycle value on design
and reverse manufacturing in semi-closed supply chain, Int. J. Production Economics 113 (2008) 528-545
5. Cooper, M.C. and Gardner, J.T., 1993. Building good business relationships: More that Partnering or strategic alliances? *International Journal of Physical Distribution and Logistics Management 23*, 6 (1993), 14-26.
6. Environmental Consequences in Reverse Manufacturing for the Computer Industry, *Journal of Cleaner Production*, 11/4 (2003): 445-458.
7. Ginter, P.M., Starling, J.M., 1978. “Reverse distribution channels for recycling,” California Management Review 20 (3), (1978) 72-82.
8. Jurg M. and Egler, Haans-Peter, 2003. “From cleaner production to sustainable Industrial production modes,” Swiss State Secretariat for Economic Affairs (SECO) Effingerstr.1, 3003, Bern, Switzerland.
9. Guide Jr., V.D.R. and Harrison, Terry P., 2003. The Challenge of Closed-loop Supply Chains. Interfaces Vol. 33, No. 6, November-December 2003.
10. Guide Jr., V.D.R., Van Wassenhove, L.N., 2002. The reverse supply chain. Harvard Business Review 80 (2), 25-26
11. Guide Jr. V.D.R., Souza, Gilvan C., Wassenhove, Luk N., Blackburn, Joseph D., 2006 Time Value of Commercial product returns, Management Science, Vol. 52, No.8, August 2006, pp.1200-1214.
12. Guide Jr., V.D.R., Jayaraman, V., Srivastava, R., Benton, W.C., 2000. Supply chain Management for recoverable manufacturing systems. Interfaces 30(3), 125-142.
13. Gland, Switzerland.|IUCN - The World Conservation Union, UNEP - United Nations Environment Programme, WWF - World Wide Fund for Nature.
14. Jayaraman V, Patterson RA, Rolland E., 2003. The design of reverse distribution Networks: models and solution procedures. European Journal of Operations research 2003; 150(1):128-49.
15. Krikke, H., van Harten, A., Schuur, P., 1999. Business case Oce: reverse logistics network Re-design for copiers. OR Spectrum 21, 381-409.
16. Kroon, L., Vrijens, G., 1995. Returnable containers: An example of reverse logistics. International Journal of Physical Distribution and Logistics Management 25 (2), 56-68.
17. Min, Hokey, Ko, Hyun Jeung, Ko, C.S., 2006. A genetic algorithmic approach to Developing the multi-echelon reverse logistics network for product returns Omega 34 (2006) 56-69
18. Mitra, Subrata, 2005. A Survey of the Third-Party Logistics (3PL) Service Providers in India, Indian Institute of Management Calcutta, WPS No. 562/ October 2005.
19. National Safety Council, “Electronic Product Recovery and Recycling Baseline Report: Recycling of Selected Electronic Products in the United States,” (Washington, DC: National Safety Council, 1999)
20. O’Rourke, D., Connelly, L., Koshland, C., 1996. “Industrial Ecology: A critical Review,” International Journal of Environment and Pollution, Vol. 6, Nos. 2/3, pp. 89-112.
21. Sarkis, J., 1995. Manufacturing strategies and environmental consciousness. Technovation 15 (2), 79-97.
22. Senge, P.M., Carstedt G., 2001. "Innovating Our Way to the Next Industrial Revolution." Sloan Management Review. Winter 2001, Volume 42, Number 2, pp. 24-38. (Beckhard Prize Winner)
23. Srivastava, S. K., 2007. Green supply chain management: a state-of-the-art literature Review. International Journal of Management Reviews 2007; 9(1):3-80.

24. Srivastava, S. K., 2008. Network design for reverse logistics, Indian Institute of Management, Lucknow 226013, India, Omega 36 (2008) 535-548
25. Stuart, Julie Ann, Low, Ming Kaan, Williams, David J., Turbini, Laura J., Ammons, Jane Chumleve, 1998. IEEE Transaction on components, packaging, and manufacturing technology – part c, vol. 21, NO.3, July 1998.
26. Thierry, Martijn, Salomon, Mark, Nunen, Jo Van, and Wassenhove, Luk Van, 1995. "Strategic Issues in Product Recovery Management," *California Management Review*, 37/2 (Winter 1995): 114-135.
27. Toffel, Michael W., 2003. "The Growing Strategic Importance of End-of-Life Product Management," *California Management Review*, 45/3 (Spring 2003): 102-129.
28. Toktay, B., Wein, L., Stefanos, Z., 2000. Inventory management of remanufacturable products. Management Science 46, 1412-1426.
29. White, C.D., Masanet, E., Rosen, C.M., and Beckman, S., 2001. "Product Recovery: An Overview of Management Challenges" MIT Sloan management review, ISSN 1532-9194, Vol. 42 No 2, pp. 24-38.

A Hindi Speech Actuated Computer Interface for Web Search

Kamlesh Sharma
Research Scholar, Dept. of CSE
Lingaya's University
Faridabad, Haryana, India

Dr. S.V.A.V. Prasad
Dean of R&D
Lingaya`s University
Faridabad, Haryana, India

Dr. T. V. Prasad
Dean of Computing Sciences
Visvodaya Technical Academy
Kavali, Andhra Pradesh, India

***Abstract*— Aiming at increasing system simplicity and flexibility, an audio evoked based system was developed by integrating simplified headphone and user-friendly software design. This paper describes a Hindi Speech Actuated Computer Interface for Web search (HSACIWS), which accepts spoken queries in Hindi language and provides the search result on the screen. This system recognizes spoken queries by large vocabulary continuous speech recognition (LVCSR), retrieves relevant document by text retrieval, and provides the search result on the Web by the integration of the Web and the voice systems. The LVCSR in this system showed enough performance levels for speech with acoustic and language models derived from a query corpus with target contents.**

Keywords- Web search; Hindi speech; HSACIWS; computer interface; human computer interaction.

I. INTRODUCTION

Information contained on the World Wide Web is inaccessible to many people. The web is primarily a visual medium that requires a keyboard and mouse to navigate, and this disenfranchises several types of users. People who lack in skills to use a keyboard and mouse find navigation difficult. Handicapped users also cannot use keyboard and mouse and people who do not have access to an Internet-capable computer have difficulty even accessing the Web. Speech recognition and generation technologies offer a potential solution to these problems by augmenting the capabilities of a web browser.

In the present development of human computer interaction (HCI), Automatic Speech Recognition interface (ASRI) is an emerging technology for offering users a totally new way of mouse, keyboard and joystick control, by using speech [1]. ASRI systems can provide both typing and web browsing for the disabled who cannot use real mouse and keyboard to obtain information from Internet as normal person do. HSACIWS users can input the query in the form of speech on the web page of the browser by using headphone. After confirmation, HSACIWS can capture the uniform resource locator (URL) and open the web page they are interested in. This novel application integrates functions of character input and cursor control, and may help those disabled to obtain information by using the search engines. World Wide Web as a repository of information for unconstrained and wide dissemination, information is now broadly available over the internet and is accessible from remote sites.

II. NEED OF INTERFACE

As network and internet have become popular and used by everyone around the word for accessing the data stored in storage mediums attached to network and servers on the internet. Data could be text, images, video, audio and other representation and non representation information. Search engines and directories were used in making queries and searches of stored data then returning the result of the query to the user. Currently used search engines and directories Google, Yahoo, Alta Vista, Ask, MSN Search, AOL, Lycos, Looks Smart and other search engine provide their search service via servers connected to networks and internet. [2]

The search engine primary mechanism is to navigate to a web page requested by user when the page is stored on any server in internet. The users manually type characters words or phrases known as a query into the search engine form. No search engine currently offers an implementation solution allowing user to make queries by speaking the query term in Hindi voice, converting the Hindi speech words into data then converting them to English word query and finally processing the query to perform a search.[3]

Speech recognition technology was developed over the past three decades. It is used in many fields like automatic speech recognition directory, military, defense, medical science, bio-informatics, home automation systems, word processing system, dictation system, embedded systems, query processing and many more systems developed for handicapped persons [7]. Most of the systems work on English language and they function at variable level of effectiveness due to limitation of the software understanding and the complexity and the variability of human speech. Large numbers of methods have been developed to acquire the accuracy in English language.

In order to enable a wider proportion of the population of India to benefit from the internet, there is a need to provide additional interface between the user and machine. Speech being a natural mode of communication among human beings can also be a convenient mode of interaction with a computer. Internationally, efforts are already on to combine hypertext navigation with spoken language. This is of particular significance in India where the rate and level of literacy are quite low. Coupled innovatively with visuals, speech and sound can add a new dimension for conveying information to the common man. It is desirable that the human-machine

interface permits communication in one's native language. This is an important issue in a multi-lingual country such as India where large numbers of people are comfortable with Hindi language. If the human oriented information over the internet can be access in Hindi language, the computer can process such hypermedia document and provide the information appropriately to a large number of users [14].

In India, Hindi speech recognition is upcoming field of research and people are working in many aspect of Hindi speech recognition. The Hindi speech engine could be utilized in search engine for querying will make the learning and use very easy for the common man. Such type of interfaces can be used by handicapped also and most importantly the person need not be aware about how to use computers.

III. System Design

The manner in which users interact with a program is known as its user interface. The user interface controls how data is entered and how information is displayed. There are mainly five types of user interfaces: character user interface (CUI), graphical user interface (GUI), touch user interface (TUI), voice user interface (VUI) and brain signal user interface (BSUI). Herein a voice user interface comes into existence. Fig. 2 shows a flowchart of Hindi speech actuated computer interface for web search shows the system design of speech actuated interface.

As shown in Fig. 1, the speech actuated computer interface consists of wearable headset and a laptop computer or desktop computer. Headphone is used to take the speech query from user. A good quality headphone acquired, amplified and digitalized the Audio signal, then wirelessly or transmitted it to the computer. The software running on laptop computer analyzed the signal and performed the instructions issued by the user. The voices utterances are of words fed to statistical speech recognition model using Hidden Markov model (HMM) where the word that were utters most likely are determined. A database was constructed with a list of word defining specific subject like fruits, vegetable, news, recipes, stock, weather etc. [9]

The uttered words are compared to the database words, if uttered word match is found a set of keywords are formed to make a query. This query is input to a search engine and the search engine processing the query and returns the result.

A user would use a speech actuated user interface in which user query is received through a device in signal form of speech and converts these speech signals to digital signals. The digital signals received as input to speech recognition module accepts the natural continuous speech patterns and generates most probable words uttered by using the HMM. The output from speech recognition module is searched against a large database of words stored in previously formulated and trained database. The output from speech recognition module is Hindi words send as input to "Speech conversion from Hindi to English" module which converts these Hindi words into corresponding English words.

The strings of English words is passed to a "Find out keywords and drop the stop constant" module for processing as in Fig. 2. Keywords are identified and marked and that recognized speech which was not useful or will not be used for search query were dropped. The keywords passed to "Make a Query" module formed a new query with the keyword and passed to "Search engine module". In case keywords are not found then the query was passed again through the user Hindi query module. The search engine searched the pages on the internet and fetched the pages on the client machine and finally displayed the results on user's screen. In case results were not found then it displayed a message on user's screen suitably.

The forming of user query was basically a combination of words that were identified from utterance using HMM. HMM [5] was used for representing speech units. Forty eight phone-like acoustic-phonetic units were used to represent Hindi sentences. Only monophone models were used due to paucity of sufficient speech data. The plosives were represented as two units-the first representing the closure, and the second representing the rest of the stop consonant. While distinction was maintained between the releases of aspirated and unaspirated plosives, no such distinction was retained between their closures. In addition to the Hindi phonemes, a few commonly occurring English vowels such as /ae/ were included. At the model level, gender identification technique can be used and the test feature vector sequence can be matched with a gender dependent HMM. Initialisation of emission probability densities of HMM states using a segmented and labelled speech database [6] should lead to better models.

Fig. 1 User interacting with system without mouse and keyboard

Dropping of stop constant and identification of key words was very important to the searching process, as shown in Fig. 2. Search engine returned results on the basis of query formed with the help of keywords identified by find keywords module.

For instance, if user gave a query for processing to the HSACIWS is "*Aaj Dili ki mandi mein aalu ka bhav kya hai*". This query was given as user query in Hindi module. The recognized string of words send as input to speech conversion from Hindi to English module which converts the Hindi string to corresponding string with the help word model and morphological analyzer is "what is the price of potatoes in the market of Delhi today". This converted string is passed as input to find the keywords and dropped the stop constant

module which tokenized the keywords like “today”, “what”, “is”, “the”, “price”, “of”, “mango”, “in”, “Delhi”, etc. and drop the stop constants. These keywords were passed as input to make a “Make a query” module which applied preprocessing techniques to generate a suitable query for search engine.

This was for HSACIWS to make sure that the input query was right as recognition was error prone. The system needed to identify the errors and came up with appropriate strategies to overcome the errors. A query was identifying in different ways as shown in Table 1. The given query “*Aaj Dili ki mandi mein Aalu ka Bhav kya hai*” could be recognized as:

Table 1. Query recognition with different variation in keywords

Aaj Bili ki mandi mein Aalu ka Bhav kya hai
Aaj Dili ki dandi mein Aalu ka Bhav kya hai
Aaj Dili ki mandi mein Balu ka Bhav kya hai
Aaj Dili ki mandi mein Balu ka Bhav kya hai
Aaj Bili ki mandi mein Balu ka Bhav kya hai
Aaj Dili ki mandi mein Aalu ka kya hai
Aaj Dili ki mandi mein ka Bhav kya hai

In such type of cases when keywords were not recognized correctly, control was passed back to user query in Hindi module and the user repeated the query again.

The results were displayed but in many cases was not sufficient to answer the user. At this point the user was allowed to optionally select to run a different module to mark hyper text links or filter information from hyper links related to the user query. Depending on the user’s choice, the results could then be displayed as text on the screen or played as speech. The output can be seen on screen in two ways, viz., (a) hyperlink number technique as shown in Fig. 3 or (b) employing a data filtering algorithm to get the actual results [10].

In hyperlink number technique a visual numbering was provided to the hyperlinks and indicating to the user how to activate them. The user can activate the hyperlink number module by providing visual number to the module [7]. Data is filtered and processed using a data filtering algorithm according to the user queries and results are sent to the speech synthesizer which translates the result from text to speech before playing it on the client system. [4]

The following are the major issues involved in the development of HSACIWS:

- *Noisy environment:* The target users of HSACIWS were primarily the farmers in rural and semi urban areas. The quality of speech signal was affected by the distance of microphone, environment where system was placed, speech codec’s and communication channel. HSACIWS system was expected to work in noisy environments, including background speech.
- *Dialect/Pronunciation variation:* The user spoke in different styles. Each dialect differed from the other at phonetic, phonological, morphological, grammatical and lexical levels. Some time moods (anger, illness, happiness and sadness) of the users effected the dialect/pronunciation variation.
- *Unstructured conversation:* The target audience of the HSACIWS may not have interacted with a computer based information access system. Hence, the conversation was typically unstructured and was filled with inconsistencies including repeats and false starts.

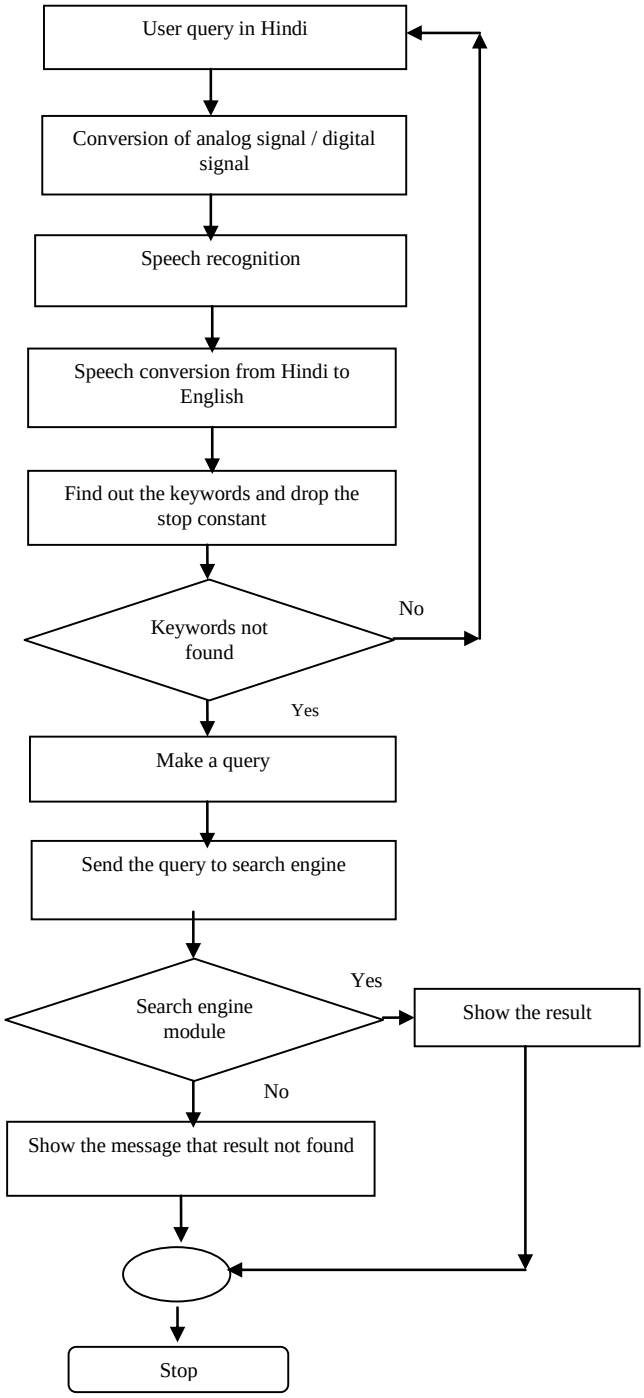


Fig. 2 Flowchart of Hindi speech actuated computer interface for web search

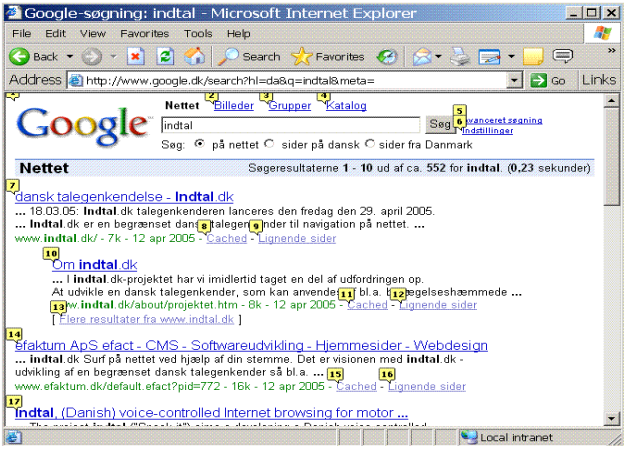


Fig. 3 Displaying hyper links with number in the web Browser [8]

A. *Training Module*

The system was trained by actual speakers articulating word continuously. Continuous speech was marked by sounds or phonemes that were connected to each other. The audio signal was processed and features extracted. The signal was smoothed by using different filters to form feature vector which were computed periodically, say every 10 to 20 milliseconds. Many types of features were used including time and frequency masking, tasking of inverse Fourier transforms resulting in a mathematical series of which the coefficient were retained as feature vector. The features were handled mathematically as vectors to simplify the training and recognition computation.

Fig. 4 describes an approach to trained the system for reading the spoken words used in the training were listed in a lexicon and a phonetic word model by using the HMM were from lexicon and phonetic spelling. These HMM word models were iteratively compared to the training speech. Training speech was produced by these HMM words models and the grammar was established with the lexicon a single probabilistic grammar for the sequence of phonemes was formed and stored in dictionary.

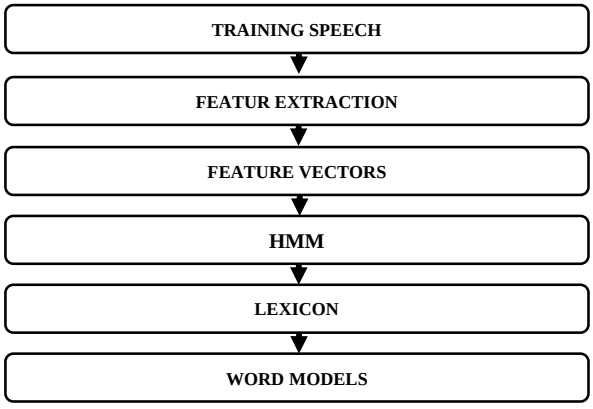


Fig. 4. Functional block diagram of the Training Module

B. *Speech Recognition Module*

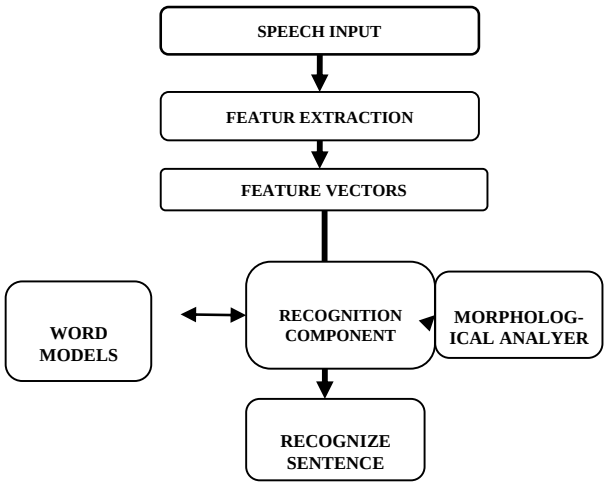


Fig. 5. Functional block diagram of the Speech Recognition Module

Recognition of an unknown speech begins with extracting feature and generating a feature vector for a particular speech by using different vectors generating technique as discussed in training module. Output of feature vector was passed to recognition component as shown in Fig. 5. Recognition component used HMM model sequences allowed by the grammar were searched to find the word sequence in the word model with the highest probability of generating that particular sequence of feature vector.

C. *Morphological Analyzer*

Morphological analyer component was used to generate a complete sentence with the help of words recognized by recognized component using the reverse and forward methodology. For translation from one language to other, the source language was first analyzed for finding the required attributes. In the source sentence the words may exist in any of their forms, so we first found their root words and then other attributes. Finding root words in the source language is called Reverse Morphology, also known as Morphological Analyzer. For the target language, the words from the given root word and their attributes were generated, and hence called forward morphology. The Morphological Analyzer constitutes of following sub modules: Input Module, Input Normalizer Module, and Tagger Module, as shown in Fig. 6 and which are described below: [13]

a) *Input Normalizer:* Input Normalizer separates the entered text into the words. It separates the words and stores them whenever blank space is encountered and provides it as input to other modules. This module also searches for the presence of the auxiliary words in the sentence and removes them, if present.

The auxiliary and its attributes are stored for further processing. The presence of auxiliary is language dependent as some languages like Sanskrit does not have auxiliaries. This module only removes the extra spaces between the words and sends the normalized sentence for further processing. [11]

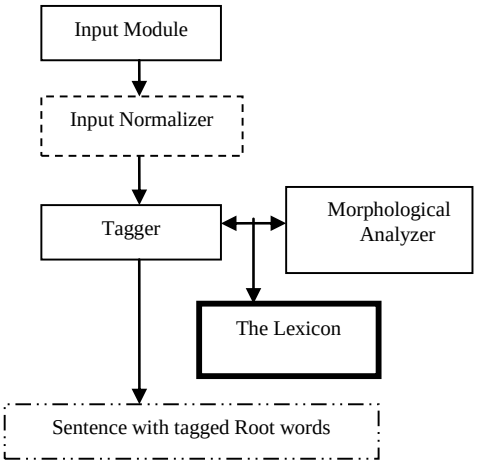


Fig. 6. Functional block diagram of the Morphological Analyzer

b) Tagger: The tagger takes the normalized sentence from the previous module and subdivided into it lexical items. Lexical items are not necessarily single words. More than one word in the input sentence may form a lexical item (e.g. give up, put off, etc). The process of dividing the sentence into lexical items is often known as Lexical Analysis. Given "What is the price of potato in Delhi market today" as input, this module would tokenize the words into an array of lexical items. Once the lexical items are obtained the next task of the tagger is to obtain the category and subcategory information for each of these items. It uses morphological analyzer and source lexicon to obtain the category and subcategory information for the lexical items.

c) Source Language Lexicon: The bilingual lexicon took Hindi as the source language and English as the target language. The lexicon contained the categorization and sub categorization information about source and target words, as in Table 2. For example:

Table 2. Categorization information

Category	Subcategory information stored in lexicon
Noun	gender, number, person, noun case
Verb	tense, aspect, gender

A given word may carry different categories and subcategories information both in the source as well as in the target language. Further, there may not be exact matching of attributes from the source language to the target language. For instance, the verb in Hindi language has two numbers (singular, plural) and two genders (masculine, feminine) whereas English language verbs does not have gender. The lexicon of a language is its vocabulary, including its words and expressions. More formally, it is a language's inventory of lexemes.

The lexicon includes the lexemes used to actualize words. Lexemes are formed according to morpho-syntactic rules and express sememes. In this sense, a lexicon organizes the mental vocabulary in a speaker's mind: First, it organizes the vocabulary of a language according to certain principles (for instance, all verbs of motion may be linked in a lexical network) and second, it contains a generative device producing (new) simple and complex words according to certain lexical rules. There were five tables a) Hindi root lexicon, b) Hindi verb feature, c) morphological lexicon, d) noun feature and e) suffices. Usually a lexicon is a container for words belonging to the same language. [12]

d) Morphological Analyzer: This module is a part of tagger which finds out category and sub-categorical information of the lexical items. The Morphological Analyzer is an integral part of any Natural Language Processing system, especially in the context of Indian languages. For fixed word order languages, the semantics of a word are primarily governed by its absolute and relative position inside a sentence. However, for free word-order languages, any clues about the semantics cannot be obtained from its position in the sentence. In case of Indian languages, which are mostly free order (like Hindi); the semantics (part of speech and other subtleties) are heavily dependent on the surface structure of the word. The task of the morphological analyzer is to identify the structural components of a word, and hence glean information about it. [13]

When an input sentence was fed for translation, the morphological analyzer identified the proper words in the sentence and retrieved necessary information about those words from the lexical database. Lexical database stored only the root form of words and its syntactic and semantic information. With the help of paradigm files, root word was extracted from the original word and all the information about that word was retried.

IV. Results And Discussions

A. Training and Testing Data Scenario for Experiments

In order to compare the effectiveness of the HSACIWS system under a scenario of truly limited data resources. There is a collection of training and testing data for the HSACIWS system that contained extremely limited amounts of data. The Training and testing data was extracted from the Indian urban and semi urban areas that have been collected through the questionnaires, Templates and personal talk. The large number of data has been collected through the questionnaires where the users were asked for different question for which the system could train and test.

The limited data consisted of the following resources:

a. *Data Corpus:* The database contains 478 word-aligned phrases and sentences from the user in urban and semi

urban areas. This elicited data collection includes both the training and testing phrases and sentences.

b. *Small Hindi-to-English Lexicon:* The database contained 2390 Hindi to English translation pairs in database. The "Speech conversion from Hindi to English" module of the system used the database and runs each Hindi input word through the morphological analyzer. The Morpher returns the root form of the word, along with a set of morphological features. The root form is then matched against the lexicon, and the set of lexical transfer rules that match the Hindi root are extracted. The morphological features of the input word are then unified with the feature constraints that appear in each of the candidate lexical transfer rules, pruning out the rules that have inconsistent features and generate the English sentence.

c. *Small English-to-Hindi Lexicon:* The database contained 2105 English to Hindi pairs in database. The "Show result" module uses the database to generate the result from English to Hindi.

B. Experimental Testing

Experiments were conducted to evaluate the baseline system and the improved HSACIWS. The HMM Toolkit was used in this experiments. A set of 100 sentences were randomly chosen from the set of 478 training sentences for test the HSACIWS. A set of 20 subjects were randomly chosen from 100 subjects (students, farmers, housewives and teachers) and were asked to raise a query to the HSACIWS to retrieve the result accordingly and show the user. The user were not trained or provided any information. Initially the experiment is done on pre define queries. The performance of the HSACIWS was calculated by using the percentage performance formula Accuracy which is defined as

$$\text{Accuracy} = \frac{S - E_s - E_d}{S} \times 100\%$$

where S, E_s, and E_d denotes the total number of sentences in the test sentences, the number of substitution errors, and deletion error respectively. The Accuracy as 100 times the ratio of the number of complete sentences recognized correctly to the total number of sentences in the test suite and show the result.

The success of each trial was based on whether the system was able to retrieve the required information to the user or not. For example user will raises the queries to the HSACIWS and system respond accordingly.

User: Sone ka Bhav kya hain.

System: Sone ka Bhav 31500 rupai hai.

User: WHO ka kya matlab hai.

System: WHO ka matlab World Health Organization.

User: Bharat ki Rajdhani kya hai.

System: Bharat ki Rajdhani delhi hai.

Over 200 experiments were conducted on the HSACIWS with 20 users (10 male and 10 female) on 100 different queries. The results are as shown in Table 3 and Table 4:

Table 3. Performance analysis of HSACIWS

Data Type	Number of Sentences	User	Sex	Total	Accuracy
Training	478	12	M/F	5736	79.3%
Test	100	10	M	1000	78.9%
Test	100	10	F	1000	77.8%

Table 4. Overall system performance analysis

System	Accuracy
HSACIWS	78.5%

The success of system was based on the confidence whether the system was able to retrieve the required information correctly to the user or not.

V. CONCLUSION

We present initial efforts for utilizing spoken Hindi language as a means of communicating web information to common Indians. We presented a voice query retrieval system in Hindi applied to document search on Internet or network.

The results of this experiment suggest that native Indians who are not able to use the computer and/or lack English skills will be able to use voice based control to navigate and obtained responses from the World Wide Web. They will not need to train the speech recognition software to their specific voice.

This system provides access to digital content over the internet to illiterate, vision-impaired, urban and semi-urban Indian people who are not able to read/write English language.

The same Speech interface can be enhanced to work for different regional languages like Punjabi, Marathi, and Telugu, etc. to enhance or to extend it for different regional languages what all is needed is the transfer or translation rules of grammar, which can be generated with great ease by using the same dataset with different target languages. Hence, there is a need to design and develop special user interfaces for accessing web information by speech of different languages.

REFERENCES

[1] Takahiro Ikeda, Shin-ya Ishikawa, Kiyokazu Miki, Fumihiro Adachi, Ryosuke Isotani, Kenji Satoh and Akitoshi Okumura, Speech-Activated Text Retrieval System for Cellular Phones with Web Browsing Capability, Proceedings of PACLIC 19, the 19th Asia-Pacific Conference on Language, Information and Computation.

[2] Michael D. Goller and Stuart E Goller, Speech Interface for search Engine".United state patent, Jun. 22 2010, shett no 1 to 4.

[3] Frederick J. Damerau and David E. Johnson, Automated set up of web-based Naturel Language Interface, U S Patent, Jun. 22 2010,

[4] M.K. Brown, Grammar Representation Requirements for voice markup Langauge, Bell Labs, Murry Hill, NJ, Dec. 1999.

[5] Honglai Xu, Tianyi Qian, Bo Hong, Xiaorong Gao, and Shangkai Gao, A Brain Actuated Human Computer Interface for Google Search, IEEE 2009.

[6] Matrin Holley and Dieter Kubesch, Speech Recognition Method for Activating a hyperlink of an Internet page, US patent 3 Aug, 2003.

[7] Bruce Moulton, Gauri Pradhan and Zenon Chaczko, Voice Operated Guidance Systems for Vision Impaired People: Investigating a User-Centered Open Source Model in International Journal of Digital Content Technology and its Applications, Volume 3, Number 4, December 2009.

[8] Edward V Porter, Voice Recognition system, U S Patent, 9 May 1989.

[9] Cameron, Hugh: " Speech at the Interface". Proceedings of the Cost249 Workshop on Speech in Telephone Networks, Ghent 2000.

[10] Hadjadj, Djamel and Dominique Burger, "BrailleSurf: An HTML Browser for visually handicapped people". In Proc. of 14th conference on "Technology and Persons with Disabilities", Los Angeles, 2000.

[11] Mihelic, France, Nikola Pavesic, Simon Dobrisek, Jerneja Gros, Bostjan Vesnicer and Janez Zibert: Homer – A Small Self Voicing Web Browser for Blind People Laboratory of Artificial Perception, Systems and Cybernetics Faculty of Electrical Engineering, University of Ljubljana, Slovenia, 2002

[12] Robin, Michael B. and Charles T. Hemphill: Considerations in Producing a Commercial Voice Browser, W3C WS on "Voice Browsers". Massachusetts, 1998.

[13] Verb Morphology for English, available at http://www.xrce.xerox.com/competencies/contentanalysis/demos/doc/mor-eng-2.html

[14] Prahallad Kishore and Black Alan, A text to speech interface for Universal Digital Library, J Zhejiang Univ Science 6A(11):1229-123.Journal of Zhejiang University, 2005

AUTHORS' PROFILE

Ms. Kamlesh Sharmsa received her masters in Computer Sc. & Engg. degree from Maharshi Dayanand University, Rohtak, India in 2009. She is currently associated with at Lingaya's University, Faridabad in the Dept. of Comp. Sc. & Engg. as Research Scholar. She has over 7 years of teaching experience at under graduate and graduate levels. Her areas of interest are artificial intelligence, operating systems, web mining, Database Management Systems, etc.

Dr. S. V. A. V. Prasad has over 30 years of experience in industry and academics. He has received his master's degree in Electronics & Communications Engg. from Andhra University, AP, India. He earned PhD from Andhra University, Waltair, Visakhapatnam, India. He was with leading research and manufacturing companies in New Delhi, India. He also taught at leading institutions like the Delhi College of Engg. (now Delhi Technological University), Delhi for many years.. He has worked as Head of the Department of Electronics & Communications Engg., Dean of Academic Affairs and as Dean of R&D and Industrial Consultancy at Lingaya's University, Faridabad. He has lectured at various forums on subjects related to electronics, communications, audio engineering, signal processing, etc. Prof. Prasad is a member of IEEE, ISTE, etc. His research interests include audio engineering, signal processing, etc.. He has large number of papers in different journals and conferences.

Dr. T. V. Prasad has over 17 years of experience in industry and academics. He received his graduate and master's degree in Computer Science from Nagarjuna University, AP, India. He was with the Bureau of Indian Standards, New Delhi for 11 years as Scientist/Deputy Director. He earned PhD from Jamia Millia Islamia University, New Delhi in the area of computer sciences/ bioinformatics. He has worked as Head of the Department of Computer Science & Engineering, Dean of R&D and Industrial Consultancy and then as Dean of Academic Affairs at Lingaya's University, Faridabad. He is with Visvodaya Technical Academy, Kavali as Dean of Computing Sciences. He has lectured at various international and national forums on subjects related to computing. Prof. Prasad is a member of IEEE, IAENG, Computer Society of India (CSI), Indian Society of Remote Sensing (ISRS) and APBioNet. His research interests include bioinformatics, artificial intelligence (natural language processing, swarm intelligence, robotics, BCI, knowledge representation and retrieval). He has over 75 papers in different journals and conferences, and also has six books and two chapters to his credit.

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Technology 6 (2012) 315 – 323

Procedia Technology

2nd International Conference on Communication, Computing & Security [ICCCS-2012]

Binarization Based Image Edge Detection Using Bacterial Foraging Algorithm

Om Prakash Verma[a,]*, Rishabh Sharma[b], Deepak Kumar[c]

[a]*Department of Information Technology, Delhi Technological University, Delhi, India*
[b]*Ericsson India Global Services Pvt. Ltd, Software Engineer, Noida, UP, India*
[c]*National Informatics Center, Delhi State, Ministry Of Communication And IT, Govt. Of India, Delhi, India*

Abstract

Bacterial Foraging Algorithm (BFA) is one of the powerful bio-inspired optimization algorithms which attempt to imitate the single and groups of *E. Coli* bacteria. In BFA algorithm, a set of bacteria try to forage towards a nutrient rich medium to get more nutrients. In this scheme, an objective function is posed as the effort or a cost incurred by the bacteria in search of food. In the present, an approach is presented for edge detection in a binarized image using bacterial foraging algorithm. .First binarization is applied to the input image to get an image matrix consisting of only the intensity values 0 and 255 of 8-bit image and then a swarm of bacteria are enthrusted on the binary image for extraction of edge information. Edges are detected by calculating the difference between intensity values of the present pixel with each of the neighboring eight pixels. Whenever the bacteria finds this intensity difference of 255 it will treat that pixel as its food and mark it as an edge pixel.

Keywords: Bacterial foraging algorithm; Edge; optimization; Binarization and heuristic.

1. Introduction

Edge detection is a terminology in image processing and computer vision, particularly in the areas of feature

* Corresponding author. Tel.: +0-991-005-0177.
E-mail address: opverma@dce.ac.in

2212-0173 © 2012 Elsevier Ltd...Selection and/or peer-review under responsibility of the Department of Computer Science & Engineering, National Institute of Technology Rourkela
doi:10.1016/j.protcy.2012.10.038

detection and feature extraction, to refer to algorithms which aim at identifying points in a digital image at which the image brightness changes sharply or more formally has discontinuities. Edge detection is a fundamental of low-level image processing and good edges are necessary for higher level of image processing [1].The edges provide important visual information since they correspond to major physical, photometrical or geometrical variations in scene object.

Edges include the most important information in the image, and can provide the information of the object's position [2]. Edges extracted from non-trivial images are often hampered by fragmentation, meaning that the edge curves are not connected, missing edge segments as well as false edges not corresponding to interesting phenomena in the image – thus complicating the subsequent task of interpreting the image data.

Traditional approaches to edge detection consists of edge detectors by Sobel [3], Prewitt [4], Kirsch [5], Frei-Chen [6], Canny [7] etc. Sobel edge operator, the Prewitt edge operator and the Robert edge operator used first-order derivative operators. In traditional approaches, each set of operation are applied for each pixels which leads to the creation of computationally expensive algorithms.

Over the last decades, various swarm intelligence based approaches have been adopted by researches in order to solve complex computational problems such as edge detection. Particle Swarm Optimization (PSO) [8] employs a swarming in which the movements of the particles are guided by the swarm's local best position as well as global best position in the required search-space. Recently Verma *et al*. [9] proposed a new optimal approach for edge detection using universal law of gravity and ant colony optimization. In this approach, the theory of universal gravity is used to calculate the heuristic function which guides the ant towards edge pixels. In 2007, Genyun Suna *et al* [10] have introduced an edge detection algorithm based on the law of universal gravity. In this approach, every image point is assumed as a celestial body, which has relationships with other neighboring image points. Recently Verma *et al*.[11] have developed an algorithm for edge detection using BFA in which direction of movement of bacteria is found using a directional probability matrix derived from ant colony optimization (ACO).

In 2002, another evolutionary bio-inspired algorithm was introduced by Passino *et al*. [12] known as bacterial foraging algorithm (BFA). BFA is the offshoot of the behavior of some species of bacteria like E. coli. There are four stages in the life cycle of bacteria namely: 1) Chemotaxis, 2) Swarming, 3) Reproduction and 4) Elimination and Dispersal. These stages in the search space generate an optimum solution to the problem of optimization.

Binarization is a process of converting a grey scale image into a black and white image. The process is often achieved by global or local thresholding. Binarization decreases the computational cost of subsequent processing compared to grey level image information.

In the proposed approach the combination of binarization and bacterial foraging algorithm is used to tackle the edge detection problem. Binarization is used to convert the input image into binary image and bacteria goes under foraging for detecting edges after being randomly placed on the binary image.

The rest of the paper is organized as the follows. The BFA is briefly reviewed in Section 2. Then, the algorithm of the proposed edge detector is presented in Section 3. The experimental results are given in Section 4 and conclusions are presented in Section 5.

2. Bacteria Foraging Technique

Bacterial foraging algorithm attempts to model the individual and group behaviour of *E.Coli* bacteria as a distributed optimization process. Foraging can be modelled as an optimization process where bacteria seek to maximize the energy obtained per unit time spent during foraging

Since its inception, BFA has been finding many important applications in real-world optimization problems from diverse domains of science and engineering. One key step in BFA is the computational chemotaxis, where a bacterium (which models a candidate solution of the optimization problem) takes steps over the foraging landscape in order to reach regions with high-nutrient content (corresponding to higher fitness).

During foraging of the real bacteria, locomotion is achieved by a set of tensile flagella. Flagella help an

E.coli bacterium to tumble or swim, which are two basic operations performed by a bacterium at the time of foraging. When they rotate the flagella in the clockwise direction, each flagellum pulls on the cell. That results in the moving of flagella independently, leading to tumbling of bacteria. In a rich medium the bacterium tumbles less number of times, whereas in a harmful place, it tumbles frequently to find a nutrient gradient. Moving the flagella in the counter clockwise direction helps the bacterium to swim at a very fast rate.

BFA mimics the four principal mechanisms observed in a real bacterial system namely; A) chemotaxis, B) swarming, C) reproduction, and D) elimination–dispersal.

A. Chemotaxis

Chemotaxis is a very important step in the bacterial foraging process. The direction of movement of bacterium is decided depending upon the rotation of the flagella. Each bacterium decides whether it should swim (move in a predefined direction) or tumble (move in an altogether different direction).

This process simulates the movement of an *E.coli* cell through swimming and tumbling via flagella. Suppose $\theta^i(j, k, l)$ represents the i^{th} bacterium at j^{th} chemotactic, k^{th} reproductive, and l^{th} elimination–dispersal step. *C(i)* indicates the size of the step taken in the random direction specified by the tumble (run length unit). Then, in computational chemotaxis, the movement of the bacterium may be represented by

$$\theta^i(j+1,k,l)=\theta^i(j,k,l)+C(i)\frac{\Delta(i)}{\sqrt{\Delta^T(i)\Delta(i)}} \tag{1}$$

where $\Delta(i)$ indicates a random number in R^2 and *C(i)* indicates the length of step size.

B. Swarming

In this step, the bacterium that has searched an optimum path, signals other bacteria so that they can together reach the desired optimum path swiftly. As each bacterium moves, it releases an attractant to signal other bacteria to swarm towards it. Also, each bacterium releases a repellent to warn other bacteria by keeping a safe distance from them.

The cell-to-cell signaling in *E. coli* swarm can be represented by the following function:

$$\begin{aligned} J_{cc}(\theta,P(j,k,l)) &= \sum_{i=1}^{S} J_{cc}(\theta,\theta^i(j,k,l)) \\ &= \sum_{i=1}^{S}\left[-d_{attractant}\exp\left(-w_{attractant}\sum_{m=1}^{p}(\theta_m-\theta_m^i)^2\right)\right] \\ &\quad + \sum_{i=1}^{S}\left[h_{repellant}\exp\left(-w_{repellant}\sum_{m=1}^{p}(\theta_m-\theta_m^i)^2\right)\right] \end{aligned} \tag{2}$$

Where $J_{cc}(\theta,P(j,k,l))$ is the objective-function value to be added to the actual objective function (to be minimized) to present a time-varying objective function.

The coefficients $d_{attractant}$, $w_{attractant}$, $h_{repellant}$, *and* $w_{repellant}$ control the strength of the cell-to-cell signaling.

C. Reproduction

After N_c chemotactic steps, a reproduction step is taken. Let N_{re} be the number of reproduction steps to be taken. Let *S* is assumed to be a positive even integer such that

$$S_r=\frac{S}{2} \tag{3}$$

be the number of population members who have had sufficient nutrients so that they will reproduce (split in two) with no mutations.

D. Elimination Dispersal

This step regulates the population of bacteria by eliminating the bacteria that land in low nutrient or noxious regions. Dispersal on the other hand locates the bacteria in new regions, which might remain unexplored with limited number of bacteria taken. Let N_{ed} be the number of elimination-dispersal events, and for each elimination-dispersal event each bacterium in the population is subjected to elimination-dispersal with probability P_{ed} .

3. The proposed image edge detection approach

In this paper, bacterial foraging technique is used as the basis of detecting edges of the input image. Edges are detected by first applying a global threshold to the input image to form a binary image. Then a group of bacteria are placed at random position on the binary image and chemotactic steps of swimming and tumbling are performed. Bacteria's foraging behavior is driven by the difference in the intensities of the neighboring pixels. Appropriate edge pixels are found by calculating the difference between intensity values of a central pixel of interest with each of its eight neighboring pixels.

Steps of the proposed approach:

A. Image binarization

The binarization process involves the assignation of pixels to either foreground or background objects by comparing their intensity values to some prescribed or automatically selected thresholds. Thresholds are applied either globally or locally. In case of global threshold, all pixels above a defined value are set to white and all the pixels below it are set to black. It is difficult to find a global threshold value for all images hence it is often calculated experimentally. In the proposed approach, binarization process is carried out by applying a global threshold on an input image so that the image matrix consists of only values 0 and 255 (Fig 1). In present case, the threshold value selected is 127.

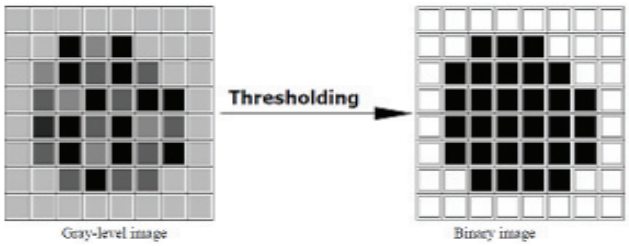


Fig 1 Conversion of gray scale image into binary

B. Edge detection

a. Initialization and search space identification

The bacteria are placed randomly all over the image which would then undergo foraging in search for food.

b. Initialize parameters

Let p be the dimension of the search space ,S be the number of bacteria placed on the image, N_c be the Number of chemotactic steps, N_{re} be the number of reproduction steps, N_{ed} be the number of elimination-dispersal steps, P_{ed} be the elimination-dispersal with probability, N_s be the Swim length and r be run-length of each bacterium.

c. Chemotactic step

i. Compute fitness function: $J(b,a,k,l)$

Fitness function of a bacterium b in its a^{th} chemotactic, k^{th} reproductive, and l^{th} elimination–dispersal step at position *(x, y)* is equal to the intensity value of the pixel at position *(x, y)*. The intensity value of a pixel will be either 0 or 255 because of binarization (Step 1).

ii. Compute cell-to-cell attractant–repellant profile to simulate the swarming behavior:

$$J(b,a,k,l)=J(b,a,k,l)+J_{cc}(\theta^b(a,k,l),\theta(a,k,l)) \quad (4)$$

Where $J(b,a,k,l)$ is the fitness function and $J_{cc}(\theta^b(a,k,l),\theta(a,k,l))$ is calculated using the Eq.2.

iii. Tumble:

A bacterium starts the process by tumbling. That is, it first moves in a random direction in search for food. The decision for this direction is done by calculating the difference between the intensity values of the present pixel (on which the bacterium is placed) with each of the eight neighboring pixels. In the above binarized image, this difference in intensity is 255 (Fig 1). Wherever the bacteria finds this intensity difference of 255, it will treat that pixel (black pixel having intensity 0) as its food and will replace this pixel with a pixel having intensity 127(it could be any random value between 0 and 255 except 0 and 255) (Fig 2).

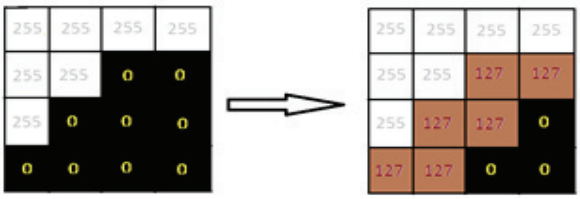


Fig 2 Allocating intensity of 127 wherever bacteria finds a difference of 255 in intensity of neighboring pixel.

These difference values are stored in a difference matrix to ease the task of finding out the pixel which would result in the maximum variation. Wherever there is an edge in an image, that corresponding pixel has a sharp increase in its value as compared to its eight neighboring pixels.

iv. Swim

The bacterium is moved to the new location using the Eq. (1). Since the bacteria aim at maximizing the energy per unit time, so they are always in search for patches with higher nutrient concentration and lesser noxious substances. It avoids noxious region and thus negative gradient region and prefers to move in positive nutrient gradient.

Thus, the bacterium follows the following criterion:

- Wherever it gets a positive nutrient gradient, it swims in that direction.
- And in neutral medium, it tumbles but if it reaches its maximum number of chemotactic steps and still not finds a positive nutrient gradient medium then it dies and it completely avoids negative nutrient gradient.

As a result, this pixel becomes the new location of the bacterium and the direction of motion of the bacterium is recorded.

d. Reproduction step

For given k and l, and each bacterium $i = 1, 2\ldots, S$

i. Sum

$$J^i_{health}=\sum_{j=1}^{Nc+1}J(b,a,k,l) \quad (5)$$

ii. Sort

Sort bacteria and in order of ascending cost J_{health}

iii. Split and Eliminate

The S_r bacteria with the highest J_{health} values die and the remaining S_r bacteria with the best values split (this process is performed by the copies that are made and are placed at the same location as their parent).

e. Eliminate and dispersal step

In the elimination-dispersal step those bacterium which are eliminated are simply dispersed to a random location. Each bacterium is subjected to elimination-dispersal with probability P_{ed}.

The Algorithm and the pseudo code

[Step 1]: Convert the input image A into a binary image B.

[Step 2]: Initialize a new matrix C which is of the same size as A, and each pixel in the C matrix is assigned the value zero (i.e, black pixel).

[Step 3]: Initialize the parameters *p*, *S*, N_s , N_c , N_{re} , N_{ed} , P_{ed} , *r*, where:

p: Dimension of the search space

S: Number of bacteria placed on the image

N_c : Number of chemotactic steps

N_{re} : Number of reproduction steps

N_{ed} : Number of elimination-dispersal steps

P_{ed} : Elimination-dispersal with probability

N_s : Swim length

r: run-length of each bacterium

[Step 4]: Elimination-dispersal loop: $l = l + 1$

[Step 5]: Reproduction loop: $k = k + 1$

[Step 6]: Chemotactic loop: $a = a + 1$

[a] For $b=1,2,...S$, take a chemotactic step for each bacterium

[b] Compute fitness function, $J(b,a,k,l)$

Let, $J(b,a,k,l)=J(b,a,k,l)+J_{cc}(\theta^b(a,k,l),\theta(a,k,l))$ (i.e., add on the cell-to-cell attractant–repellant profile to simulate the swarming behavior)

[c] Let $J_{last}=J(b,a,k,l)$ to save this value, since we may find a better cost via a run.

[d] Tumble: The direction of the next move of the bacterium is found out by finding the difference of the current pixel with each of the eight neighboring pixels.

If the value of the current pixel is 255, and as soon as it is found out that there exists a neighboring pixel whose value is zero, then the value of this corresponding pixel in the C matrix is assigned the value 127.Similarly, if the value of the current pixel is zero and if there exists a neighboring pixel whose value is 255, then the current pixel *(i,j)* in matrix C is assigned the value 127.

In either case, the bacterium is moved to this new location and the direction of motion is recorded for swimming in that direction.

[e] Swim:

1) Let $m=0$ (counter for swim length)

2) **While** ($m < N_s$)

Let $m = m+1$

3) **if** ($J(b,a+1,k,l) < J_{last}$)

Set $J_{last} = J(b, a+1, k, l)$

4) Let $\theta^b(a+1,k,l) = \theta^b(a,k,l) + C(b)\frac{\Delta(b)}{\sqrt{\Delta^T(b)\Delta(b)}}$

And use $\theta^b(a+1,k,l)$ to compute the next $J(b, a+1, k, l)$
Else let *m*= N_s
5) **End While** Loop
This results is a step of size *C(b)* in the direction of the tumble for bacterium *b*.
If j< N_c go to [Step 6][e] as the life of the bacteria is not over

[Step 7]: Reproduction:
[a] For the given *k* and *l*, and for each *i* = 1, 2, . . . , *S*, let $J^i_{health} = \sum_{j=1}^{Nc+1} J(b,a,k,l)$ be the health of the bacterium *i* (a measure of how many nutrients it got over its lifetime and how successful it was at avoiding noxious substances). Sort bacteria and chemotactic parameters *C(b)* in order of ascending cost J_{health} (higher cost means lower health).
[b] The bacteria with the highest J_{health} values die, and the remaining S_r bacteria with the best values split (this process is performed by the copies that are made are placed at the same location as their parent).

[Step 8]: If $k < N_{re}$, go to step 3. In this case, we have not reached the number of specified reproduction steps, so we start the next generation of the chemotactic loop.

[Step 9]: Elimination–dispersal: For *i = 1, 2, . . . , S* with Probability P_{ed}, eliminate and disperse each bacterium (this keeps the number of bacteria in the population constant). To do this, if a bacterium is eliminated, simply disperse another one to a random location on the optimization domain. If *l* < N_{ed}, then go to step 2;otherwise, end.

4. Experimental Results

The Experiments are conducted to evaluate the performance of the proposed approach using two test images, Camera and Einstein (Fig 3 and 4) and the performance is evaluated by comparing the edge map with conventional approaches.

Furthermore, various parameters of the proposed approach are set as follows:

$$\mu = 1, S = 100, d_{att} = 0.1, w_{att} = 0.2, h_{rep} = 0.1,$$
$$w_{rep} = 10, N_s = 10, P_{ed} = 0.95, N_{ed} = 100$$
$$N_{re} = 1, N_c = 100$$

In these experiments, traditional edge detectors are executed by MATLAB toolbox. The codes for our method were also written in MATLAB. The path traversed by a bacterium represents the edge pixels.

The proposed approach leads to less computational cost during the edge detection process because of binarization as compared to traditional edge detectors as they use grey level image information. The proposed method gives excellent results and robustly draws edges even in complex images. By applying global thresholding during binarization there are some points in the background image that are considered to be foreground and vice versa thus introducing noise in the binary image. The proposed approach is not able to differentiate between edge pixels and noise. However, the results also depend upon the kind of image one is

considering for edge detection. But overall the connectivity and number of pixels is high in comparison to traditional edge detectors.

The proposed approach tends to find meaningful edges in most images and is successful in optimally solving the edge detection problem by taking into consideration both the space and the time complexity.

5. Conclusions

In this paper, a new BFO based image edge detection approach has been successfully developed. Edge detection is a critical topic in computer vision and image processing. In the proposed approach the edge detection problem is handled by first converting the input image into binary image and then running the modified bacterial foraging algorithm on that image. Input image is first converted to binary image by applying a global threshold on the image and then bacteria are distributed randomly on the binary image for foraging. Edges are detected by bacteria by calculating the difference between intensity value of the present pixel with each of the neighboring eight pixel.

References

[1] M.B. Ahmad and T.S. Choi , "Local Threshold and Boolean Function Based Edge Detection", IEEE Transactions on Consumer Electronics, Vol. 45, No 3,pp. 674 -679, August 1999.

[2] RongWang, Li-qun gao, Shu Yang, Yan-Chun Liu, "An Edge detection method by combining fuzzy logic and neural networks", Machine Learning and Cybernetics,vol. 7,pp. 4539-4543,2005

[3] SOBEL, I., "Camera Models and Perception", Ph.D. thesis, Stanford University, Stanford, CA, 1970.

[4] J.M.S. Prewitt, "Object enhancement and extraction", in: B.S. Lipkin, A. Rosenfeld (Eds.), Picture Analysis and Psychopictorics, Academic Press, New York, NY, pp. 75-149, 1970.

[5] Kirsch, R., "Computer determination of the constituent structure of biological images", Computers and Biomedical Research, Vol. 4, pp. 315-328, 1971

[6] Frei and Chen, "Fast boundary detection: A generalization and a new algorithm", IEEE Trans. Computers, Vol. 26, pp. 988-998, 1977.

[7] J. Canny, "Computational Approach to Edge Detection", IEEE Transaction on Pattern Analysis and Machine intelligence, No. 6, pp. 679-698, 1986.

[8] Setayesh M., Zang M. and Jonhnton M., "A new homogeneity-based approach to edge detection using PSO", 24th International Conference Image and Vision Computing, New Zealand (IVCNZ), pp. 231-236, 2009.

[9] Verma O.P., Sharma R., "An optimal edge detection using universal law of gravity and Ant colony algorithm", World congress on communication and information technologies (WICT), 2011, in press

[10] Genyun Suna, Qinhuo Liua, Qiang Liua, Changyuan Jib, Xiaowen Lia ,"A novel approach for edge detection based on the theory of universal gravity" Pattern Recognition, Vol. 40(10), pp . 2766-2775, 2007

[11] Verma O.P., Hanmandlu M., Kumar P., and Jindal A, "A Novel Bacterial Foraging Technique for Edge Detection", Pattern Recognition Letters, No.8, pp.1187-1196, 2011.

[12] Passino,K. M., "Biomimicary of Bacterial Foraging for distributed optimization control", IEEE control system magazine, pp. 52-67,2002.

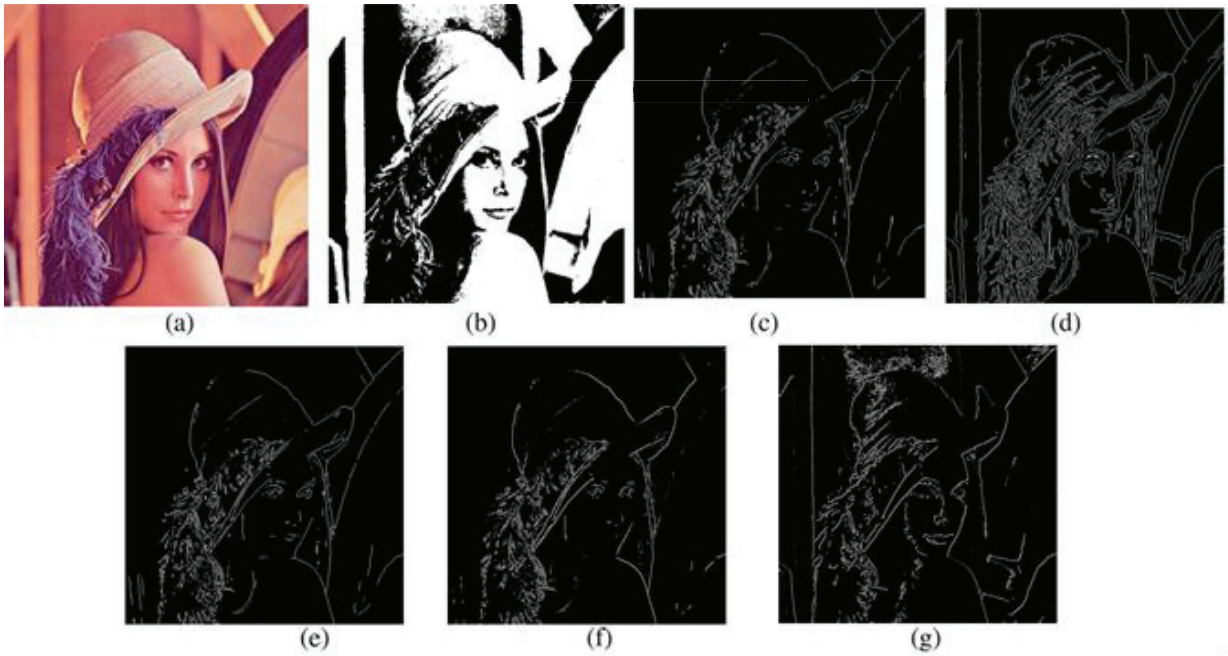


Fig. 3. (a) Original lena image (b)Binary image (c) Sobel Edge Detector (d) Canny Edge Detector (e) Prewitt Edge Detector (f) Robert Edge Detector (g) The proposed approach

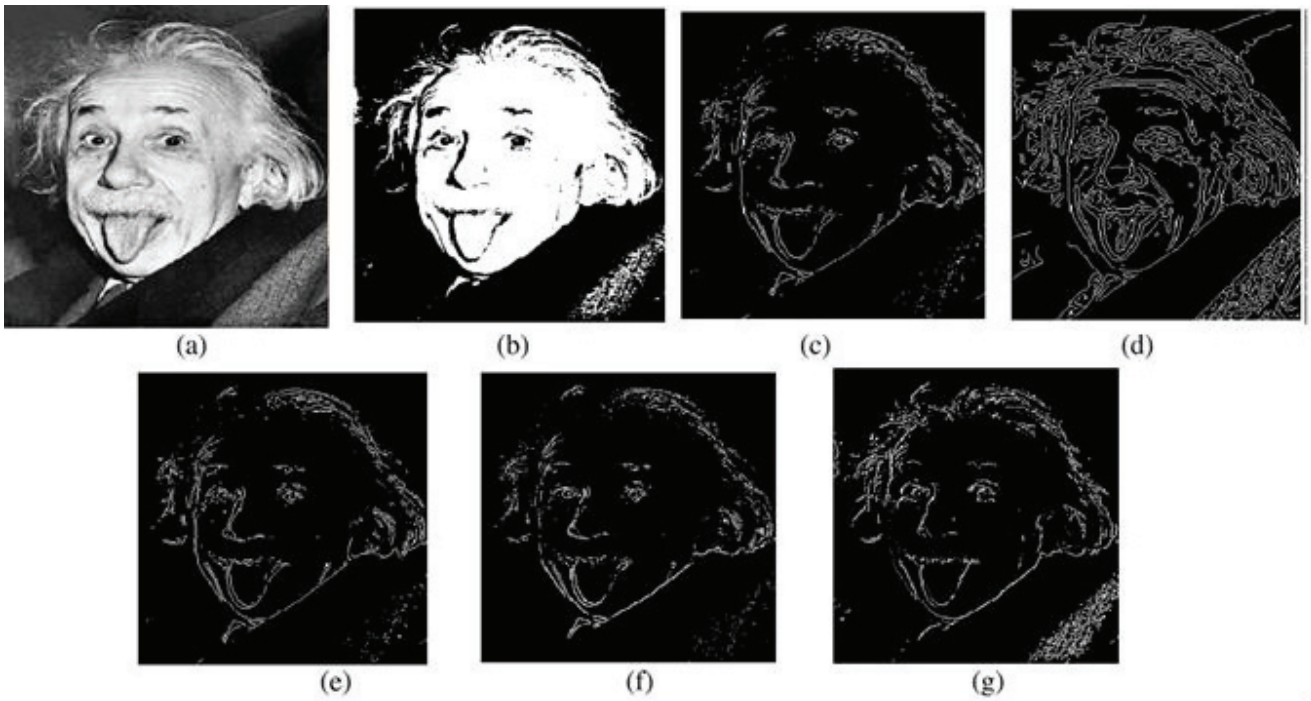


Fig. 4. (a) Original Einstein image (b)Binary image (c) Sobel Edge Detector (d) Canny Edge Detector (e) Prewitt Edge Detector (f) Robert Edge Detector (g) The proposed approach

International Journal of Pure and Applied Mathematics
Volume 75 No. 4 2012, 427-439
ISSN: 1311-8080 (printed version)
url: http://www.ijpam.eu

ijpam.eu

BOUNDARY CHARACTERISTIC ORTHOGONAL POLYNOMIALS FOR SINGULARLY PERTURBED PROBLEMS

Vivek Kumar

Department of Applied Mathematics
Delhi Technological University
Bawana Road, Delhi, 110042, INDIA

Abstract: In this paper we generate fitted mesh using the boundary characteristic orthogonal polynomials for numerically solving singularly perturbed boundary value problems. The method is based upon the orthogonal collocation method. Boundary characteristic orthogonal polynomials (BCOPs) are generated using the Gram - Schmit process from a set of linearly independent functions which also satisfy the given boundary conditions. The procedure is illustrated by taking several examples. After obtaining the fitted mesh, the problems have been solved using upwind finite difference method.

AMS Subject Classification: 65L11, 65L10, 65L50, 65M50
Key Words: boundary characteristic orthogonal polynomials (BCOPs), singularly perturbed boundary values problems (SPPs), fitted mesh methods, upwind finite difference

1. Introduction

Orthogonal polynomials have been extensively used in numerical approximations, for example, the famous Legendre, Chebyshev polynomials and many more. The importance of orthogonal projection and orthogonal decomposition,

Received: August 6, 2011

particularly in the solution of systems of linear equations and in the least square data fitting is also well known. Now a large number of books and research papers are available on orthogonal polynomials and their applications and some good references can be found in [1, 2, 3].

Problems in which a small perturbation parameter, say ϵ is multiplied to the highest derivative arise in various fields of science and engineering, for instance fluid mechanics, elasticity, hydrodynamics, etc. The main concern with such problems is the rapid growth or decay of the solution in one or more narrow "layer region(s)". These kinds of problems are known in the literature as singularly perturbed problems (SPPs).

Singular perturbation problems in consideration have shocks as boundary layers or interior layers. For such kinds of problems the solution can be smooth in most of the solution domain with small area where the solution changes very quickly. To approximate their solution it is well known (see [4, 5]) that the classical numerical methods cannot be used on uniform meshes; the reason is that the error is unbounded [6] for arbitrary values of the singularly perturbed parameter, ϵ. So when solving such problems numerically, one would like to adjust the discretization to the solution. In terms of mesh generation, we want to have many points in the area where the solution has strong variations and a few points in the area where the solution has weak variations and such method is known as fitted mesh methods.

In orthogonal collocation, we first generate the orthogonal polynomials and then we find the roots of those orthogonal polynomials (Each polynomial in an orthogonal sequence has all n of its roots real, distinct, and strictly inside the interval of orthogonality) and treat them as collocation points. The choice of collocation points is also critical and should not be arbitrary in realistic problems. The dependence of the roots according to the mesh requirement lies on the choice of the weight functions used in defining the inner product. In Section 2 we discuss how to find the BCOPs and the corresponding roots and numerical results have been presented in Section 3.

2. Generation of Mesh using BCOPs

Let us first define the inner product in the functional space for two functions $f(x)$ and $g(x)$ defined over the domain $D \in R^n$ by

$$< f, g >= \int_D w(x) f(x) g(x) dD \tag{1}$$

where $w(x)$ is the suitable chosen weight function according to the mesh defined over D. The induced norm of a function using above inner product is, therefore, given as

$$||f||^2 = \int_D w(x) f^2(x) dD. \tag{2}$$

To generate an orthogonal sequence, we can start with the set

$$\{h(x)f_i(x)\}, i = 0, 1, 2, 3.... \tag{3}$$

where $h(x)$ is the chosen function which satisfy the given boundary conditions of a differential equation and $f_i(x)$ are the linearly independent functions over the domain D. Note that each $h(x)f_i(x)$ will, therefore, also satisfy the same boundary conditions (if we have zero boundary conditions). Otherwise for non zero boundary conditions, we have to choose $f_i(x)$ which also satisfy the boundary conditions and independency.
To generate an orthogonal sequence ϕ_i, we apply the well known Gram - Schmidt process, which is given as

$$\phi_1 = hf_1 \tag{4}$$

$$\phi_i = hf_i - \sum_{j=1}^{i-1} c_{ij}\phi_j, \quad i = 2, 3, 4.... \tag{5}$$

where

$$c_{ij} =< hf_i, \phi_j > / < \phi_j, \phi_j > \tag{6}$$

The orthogonal sequence can also be normalized by dividing each ϕ_i by its norm.

2.1. BCOPs Approximations

First we try to approximate a given function, using the method of least square, which has boundary layers on both the boundaries by generating corresponding BCOPs. Suppose we take original function $F(x)$ and write it as a linear combination of generated BCOPs ϕ_is as

$$F(x) = \sum_{j=0}^{N} d_j \phi_j(x). \tag{7}$$

Then the d_js can be calculated using the concept of orthogonality (Fourier-Legendre type).

$$d_j =< F, \phi_j > / < \phi_j, \phi_j > . \tag{8}$$

Suppose we start with the function

$$F(x) = \frac{\exp((x-1)/\sqrt{\epsilon}) + \exp(-x/\sqrt{\epsilon})}{(1+\exp(-1/\sqrt{\epsilon}))} - \cos(\pi x)^2; \;\; x \in (0,1), \tag{9}$$

which has boundary layers on both the sides of the interval for small values of parameter ϵ with zero boundary conditions. Let us we start with $h(x) = x(1-x)$, as it satisfies the zero boundary conditions. The inner product for this function can be defined as

$$< f, g >= \int_0^1 \frac{f(x)g(x)}{\sqrt{x(1-x)}} dx, \tag{10}$$

with weight function $w(x) = 1/\sqrt{x(1-x)}$, as it will be very helpful for generating finer mesh at the boundary points 0 and 1. As we know that the computation of the integral involved will become simpler if we deal with the polynomial functions f_is. So we start with one of the obvious choice of f_i as

$$f_i = \{1, x, x^2, x^3, ...\} \tag{11}$$

It is clear that the function hf_i also satisfies the zero boundary conditions. Using the above procedure to generate BCOPs, Some ϕ_is can be given as

$$\phi_0 = \frac{8x(1-x)\sqrt{6}}{3\sqrt{\pi}}; \tag{12}$$

$$\phi_1 = \frac{32(x^2(1-x) - (1/2)x(1-x))}{\sqrt{\pi}}; \tag{13}$$

and so on. Similarly ϕ_9 can also be given as

$$\begin{aligned}\phi_9 = &-2.330769986x(x-1)(2x-1)(3.27680 \times 10^5 x^8 - 1.310720 \times 10^6 x^7 \\ &+ 2.146304 \times 10^6 x^6 - 1.851392 \times 10^6 x^5 + 9.02144 \times 10^5 x^4 \\ &- 2.47808 \times 10^5 x^3 + 36256x^2 - 2464x + 55).\end{aligned} \tag{14}$$

For the above $F(x)$, equation (7) can be written as

$$F(x) = \sum_{j=0}^{9} d_j \phi_j(x) \tag{15}$$

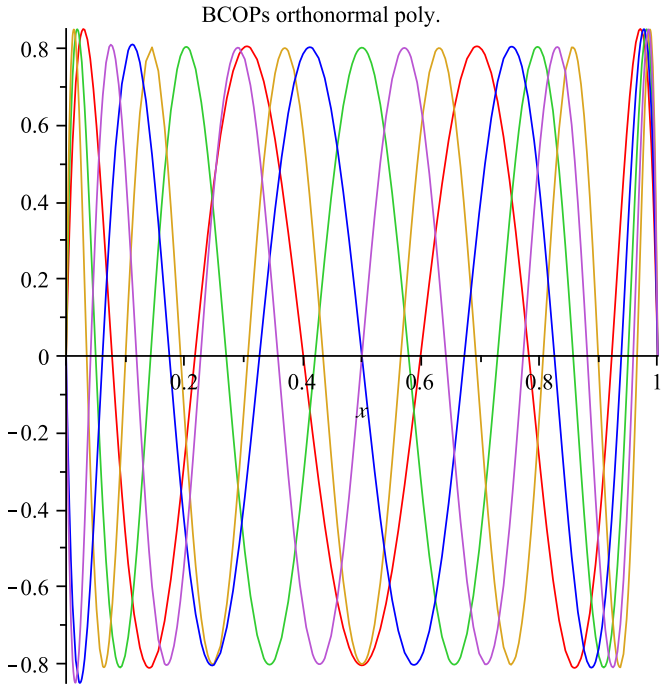


Figure 1: BCOPs for various ϕ_is

where ϕ_is as generated above. The corresponding d_is are found using orthogonal property as

$$\begin{aligned} &d_0 = -.556648, d_1 = -2.246921 \times 10^{-10}, \\ &d_2 = -.756039, ...d_{,8} = -.025290, d_9 = -1.16538 \times 10^{-13}. \end{aligned} \tag{16}$$

All the ϕ_is as used in equation (15) has been plotted in the Figure 1 and the corresponding approximated $F(x)$ is plotted in the Figure 2. As we can see in the Figure 2 that both the functions are coinciding completely with each other.

3. Numerical Results

Now we generate the fitted mesh for a given problem using the BCOPs method and solve the singularly perturbed problems using upwind finite difference methods on non-uniform meshes.

Test Problem 1. (see [7]) We consider 1D linear reaction-diffusion problem as

$$-\epsilon u''(x) + u(x) = -\cos^2(\pi x) - 2\epsilon\pi^2 \cos(2\pi x), x \in (0,1), 0 < \epsilon << 1, \tag{17}$$

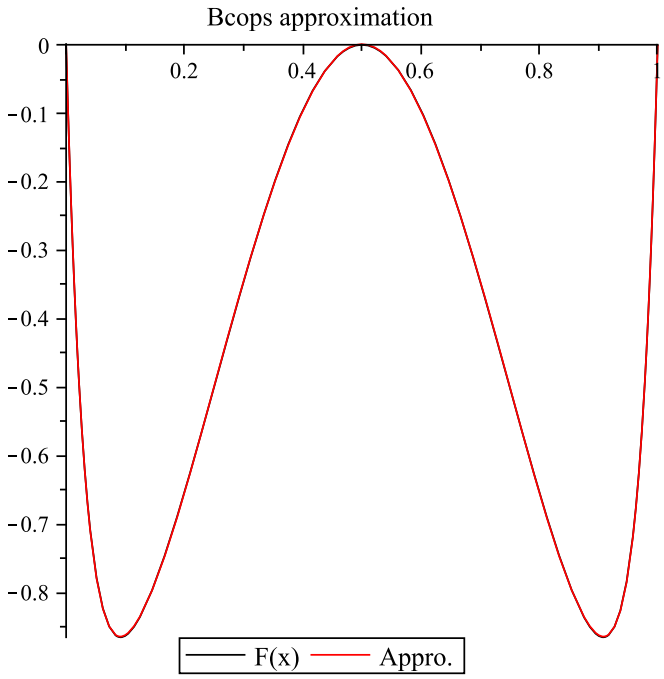


Figure 2: Approximation for $\epsilon = .001$

with boundary conditions

$$u(0) = 0,\ u(1) = 0 \tag{18}$$

whose exact solution (as also discussed in equation (9)) is

$$u(x) = \frac{\exp(-1(1-x)/\sqrt{\epsilon}) + \exp(-x/\sqrt{\epsilon})}{1 + \exp(-1/\sqrt{\epsilon})} - \cos^2(\pi x).$$

This problem has regular boundary layers of width $O(\sqrt{\epsilon})$ at $x = 0$ and $x = 1$. Using the same inner product as defined above in equation (10) we get the same orthonormal polynomials as given above equations (12 -14), we find the roots of such polynomials and apply orthogonal collocation method to compute the numerical solution. The roots of few BCOPs other than $0, 1$ are given in Table 1.

In Table 2, we have shown the maximum error for various values of perturbation parameter ϵ by using different BCOPs. Results looks better even for smaller values of ϵ. It is also clear from the Table that for each ϕ_is (say ϕ_5) maximum error is reducing as for higher degree BCOPs, the more roots are falling in the boundary layer region.

Similarly we can solve another type of problem i.e convection diffusion equation which has only one boundary layer.

$\phi_i \downarrow$	$roots \rightarrow$									
ϕ_5	.50000	.10089	.89910	.27853	.721460					
ϕ_6	.077624	.92237	.21780	.78219	.40090	.59909				
ϕ_8	.049965	.95003	.14286	.85713	.27071	.72928	.42099	.57900		
ϕ_9	.50000	.041367	.95863	.11897	.88102	.22742	.77257	.35795	.64204	
ϕ_{10}	.034806	.96519	.10054	.89945	.19348	.80651	.30731	.69268	.43428	.56571

Table 1: Roots for the corresponding ϕ_is for test problem 1

$\epsilon \downarrow \phi_i \rightarrow$	ϕ_5	ϕ_6	ϕ_7	ϕ_8	ϕ_9	ϕ_{10}
0.1	0.0706	0.0555	0.0457	0.0384	0.0330	0.0287
0.01	0.0501	0.0449	0.0392	0.0341	0.0299	0.0265
0.001	0.0288	0.0179	0.0049	0.0144	0.0244	0.0297
0.0001	0.0095	0.0156	0.0224	0.0283	0.0314	0.0308
0.00001	0.000964	0.0016	0.0026	0.0039	0.0057	0.0080

Table 2: Maximum error for various values of ϵ and mesh points for test problem 1

Test Problem 2. (see [5]) Convection diffusion equation in 1D can be given as

$$\epsilon u''(x) + 2u'(x) = 0, x \in (0,1), 0 < \epsilon << 1, \tag{19}$$

with boundary conditions

$$u(0) = 1,\ u(1) = 0 \tag{20}$$

This equation has boundary layer of order $O(\epsilon)$ at $x = 0$ and its exact solution is given as

$$u(x) = \frac{e^{-2x/\epsilon} - e^{-2/\epsilon}}{1 - e^{-2/\epsilon}}. \tag{21}$$

To find the fitted mesh (finer mesh at $x = 0$) we define the inner product as

$$< f, g >= \int_0^1 \frac{f(x)g(x)}{\sqrt{x}} dx, \tag{22}$$

with weight function as $w(x) = 1/\sqrt{x}$ to have finer mesh near $x = 0$. In this case we choose f_is as

$$f_i = \{1 - x, (1 - x)^2, (1 - x)^3, (1 - x)^4, ...\} \tag{23}$$

and $h(x) = 1 - x$. We can see that each hf_i satisfies the non-zero boundary conditions as per the requirement. Using the above procedure to generate BCOPs, we give some ϕ_is as mentioned below

$$\begin{aligned}
\phi_0 &= \frac{\sqrt{15}(1-x)}{4};\\
\phi_1 &= -\frac{3\sqrt{5}(-7(1-x)^2+6(1-x))}{8};\\
\phi_5 &= (.26757(x-1))(7429x^5 - 14535x^4 + 9690x^3 - 2550x^2 + 225x - 3);\\
\phi_{10} &= -(0.00019037(x-1))(1.14353210^{10}x^{10} - 5.05281610^{10}x^9\\
&\quad + 9.42781510^{10}x^8 - 9.66955410^{10}x^7 + 5.945469010^{10}x^6 - 2.24229110^{10}x^5\\
&\quad + 5.09611610^9x^4 - 6.5756310^8x^3 + 4.251410^7x^2 - 1.04975010^6x + 4199);
\end{aligned} \tag{24}$$

and so on. Similarly we can generate much higher order BCOPs. For this example we have generated the functions upto ϕ_{14}. The roots of the corresponding BCOPs (since $x = 0$ (left boundary point) is not the root for any ϕ_is because of our choice of BCOPs, we add $x = 0$ as an additional root for all the ϕ_is.) other than $0, 1$ are given in Table 3.

$\phi_i \downarrow$	$roots \rightarrow$									
ϕ_{10}	.00490	.04354	.11785	.22204	.34806	.48615	.62560	.75561	.86614	.94867
ϕ_{11}	.00413	.03677	.09991	.18943	.29947	.42284	.55146	.67691	.79103	.88629
	.95661									
ϕ_{14}	.00266	.02377	.06511	.12493	.20070	.28920	.38671	.48910	.59195	.69146
	.78150	.86273	.92468	.97220						

Table 3: Roots for the corresponding ϕ_is for test problem 2

Table 4 discuss the maximum error for the test problem 2 for various values of ϵ verses various ϕ_is. In this case too maximum error reduces as the value of ϵ becomes smaller.

Test Problem 3. (see [8]) We consider another interior layer SPP problem as

$$\epsilon u''(x) + xu'(x) = -\epsilon\pi^2\cos(\pi x) - \pi x\sin(\pi x), x \in (-1,1), 0 < \epsilon << 1, \tag{25}$$

with boundary conditions

$$u(-1) = -2,\ u(1) = 0. \tag{26}$$

$\epsilon \downarrow \phi_i \rightarrow$	ϕ_5	ϕ_{10}	ϕ_{11}	ϕ_{12}	ϕ_{13}	ϕ_{14}
0.1	0.2421	0.2333	0.2285	0.2238	0.2194	0.2152
0.01	0.3107	0.2730	0.2572	0.2463	0.2405	0 .2391
0.001	0.0520	0.1525	0 .1757	0 .1990	0 .2218	0.2435
0.0001	0.0055	0.0178	0.0210	0.0245	0.0283	0.0323
0.00001	0.00054	0.0018	0 .0021	0.0025	0.0029	0.0033

Table 4: Maximum error for various values of ϵ and mesh points for test problem 2

Its exact solution is not known and for small ϵ it gives a turning point near $x = 0$. Therefore we need finer mesh near the turning point to resolve the interior layer for small ϵ. To find the fitted mesh (finer mesh at $x = 0$) we define the inner product as

$$< f, g >= \int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1-x^2}} dx, \tag{27}$$

with weight function as $w(x) = 1/\sqrt{1-x^2}$ to have finer mesh near $x = 0$. In this case we choose f_is as

$$f_i = \{x, x^2, x^3, x^4, ...\} \tag{28}$$

and $h(x) = (x - 1)$. We can see that in this case $h(x)f_i$s do not satisfy the boundary conditions (i.e non zero b.c.) but at the point $x = 0$, where the interior layer occur, all the $h(x)f_i$s satisfy the given zero conditions. The roots of the few corresponding BCOPs other than 1 are given in Table 5 below.

$\phi_i \downarrow$	$roots \rightarrow$								
ϕ_7	.20449	.56782	.84597	-.18876	-.55207	-.83014	-.98063		
ϕ_8	.00620	.35035	.65300	.87768	-.33793	-.64057	-.86519	-.98471	
ϕ_9	.16224	.46129	.71569	.90057	-.15219	-.45124	-.70563	-.89045	-.98762
ϕ_{10}	.00414	.28704	.54703	.76304	.91760	-.27875	-.53873	-.75473	-.90925 -.98977

Table 5: Roots for the corresponding ϕ_is for test problem 3

Figures 3 and 4 show the results for different values of ϵ and for $\epsilon = .000001$, we can see a thin interior layer at $x = 0$.

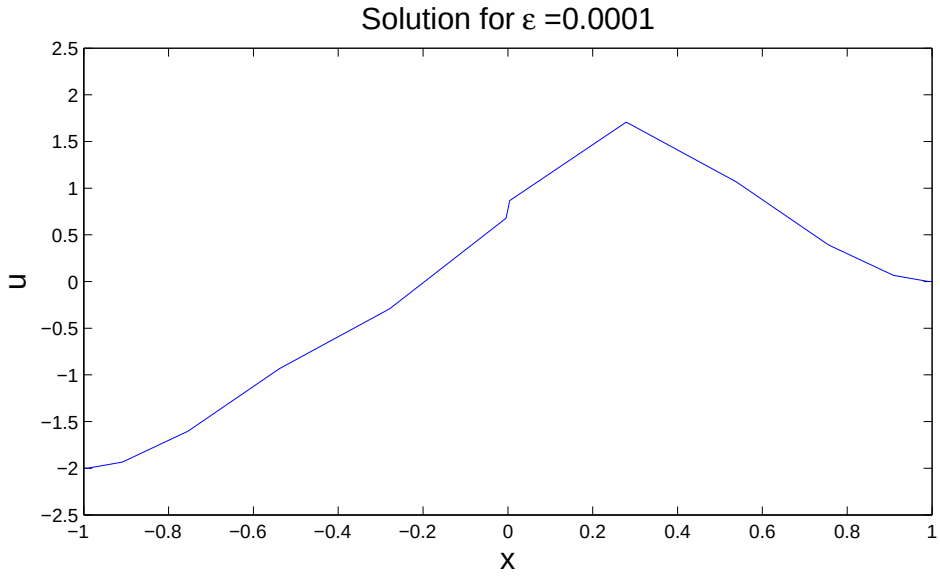


Figure 3: Approximate solution for $\epsilon = .0001$

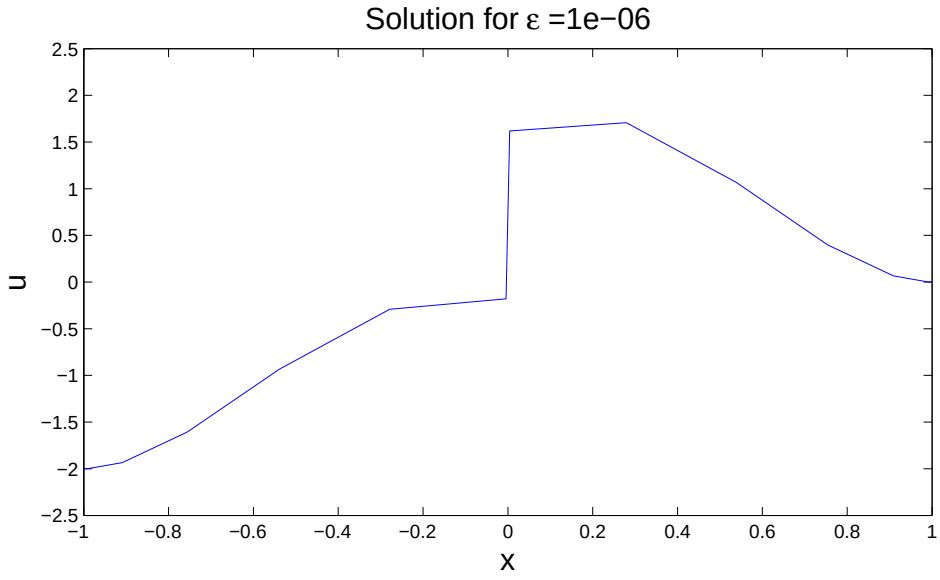


Figure 4: Approximate solution for $\epsilon = .000001$

Test Problem 4. (see [9]) We consider 2D linear reaction-diffusion problem as

$$-\epsilon^2 \triangle u(x,y) + 2u(x,y) = f(x,y), \text{ in } \Omega = (0,1) \times (0,1). \tag{29}$$

f has been chosen such that the exact solution of Eq. (29) is given as

$u(x,y) =$

$\phi_i \downarrow \epsilon \rightarrow$	2^{-5}	2^{-6}	2^{-7}	2^{-8}	2^{-9}	2^{-10}	2^{-15}
ϕ_8	0.00700	0.02057	0.01503	0.00424	0.00106	0.00026	0.00000026
ϕ_{10}	0.01570	0.00685	0.02144	0.00848	0.00217	0.00054	0.00000053

Table 6: Maximum error for various values of ϵ and mesh points for test problem 4

$$\left(1 - \frac{\exp(-x/\epsilon) + \exp(-(1-x)/\epsilon)}{1 + \exp(-1/\epsilon)}\right)\left(1 - \frac{\exp(-y/\epsilon) + \exp(-(1-y)/\epsilon)}{1 + \exp(-1/\epsilon)}\right).$$

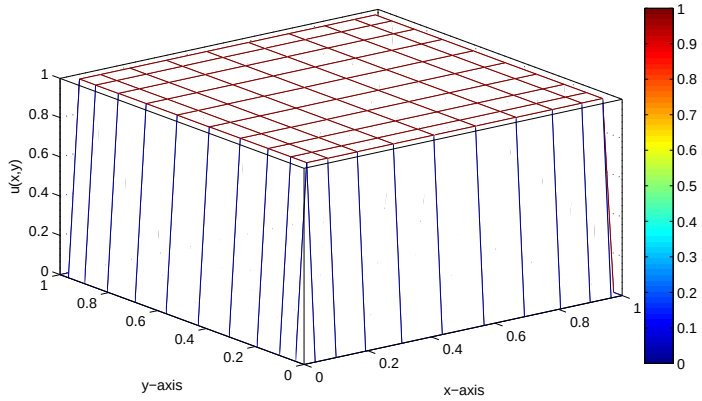
This $u(x, y)$ has typical boundary layers of width $O(\epsilon)$. Since the exact solution is known, we can accurately measure the maximum error as given in the Table 6. We have taken the same ϕ_is as discussed in example 1 (with zero boundary conditions). To solve 2D problem we have used the same 1D derivative matrices to generate Laplacian using tensor products, also known as *Kronecker products* as discussed in [10]. Table 6 gives the maximum errors, and Figure 5 shows the exact and computed results.

4. Conclusion

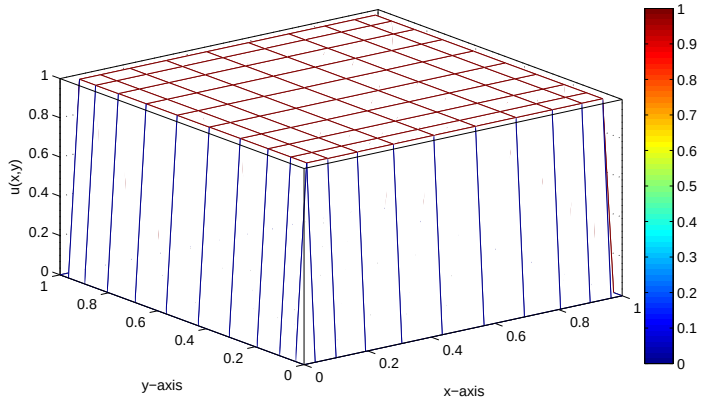
In this paper we have introduced the BCOPs methods to study the singularly perturbed problems having boundary and interior layers in one and two dimensions. The one of the advantage of BCOPs is that all the corresponding polynomials also satisfy the boundary conditions and there is no need to satisfy them separately. BCOPs have been used to solve problem in the vibration by treating them as the bases functions for the approximation. But in this paper we have taken the advantage of orthogonal collocation to develop fitted meshes to deal with SPPs.

Acknowledgement

Author would like to acknowledge Department of Science and Technology (DST), New Delhi, INDIA, for the financial grant under the fast track project SR/FTP/ MS-04/2008 and Prof. Bani Singh for all kind of help to learn and impliment BCOPs.



(a) Exact sol. for $\epsilon = 2^{-15}$ using ϕ_{10}



(b) Computed sol. for $\epsilon = 2^{-15}$ using ϕ_{10}

Figure 5: Solutions for example 4 (29)

References

[1] Bani Singh, S. Chakraverty, Boundary characteristic orthogonal polynomials in numerical approximation, *Communications in Numerical Methods in Engineering*, **10** (1994), 1027-1043.

[2] Bani Singh, S. Chakraverty, Flexural vibration of skew plates using boundary characteristic orthogonal polynomials in two variables, *Journal of Sound and Vibration*, **173** (1994), 158-178.

[3] Walter Gautshi, *Orthogonal Polynomials: Computation and Approximation*, Numerical Mathematics and Scientific Computation, Oxford University Press, USA (2004).

[4] J.J.H. Miller, E. O'Riordan, I.G. Shishkin, *Fitted Numerical Methods for Singular Perturbation Problems*, World Scientific (1996).

[5] P.A. Farrell, A.F. Hegarty, J.J.H. Miller, E. O'Riordan, I.G. Shishkin, *Robust Computational Techniques for Boundary Layers*, Chapman and Hall, CRC (2000).

[6] Vivek Kumar, High order compact finite difference scheme for sing. pert. reaction diffusion problems on a new mesh of Shishkin type, *J. of Optim. Th. and Appl.*, **143** (2009), 123-147.

[7] E.P. Doolan, J.J.H. Miller, W.H.A. Schilders, *Uniform Numerical Methods for Problems with Initial and Boundary Layers*, Boole Press, Dublin (1980).

[8] F. Mazzia, A. Sestini, D. Trigiante, The continuous extension of the B-spline linear multistep methods for BVPs on non-uniform meshes, *Applied Num. Math.*, **59** (2009), 723-738.

[9] J. Li, I.M. Navon, Uniformly converg. finite element methods for sing. pert. elliptic boundary value problems I: Reaction-diffusion type, *Comp. Math. Appl.*, **35**, No. 3 (1998), 57-70.

[10] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Press, Cambridge, UK (1991).

ARTICLE IN PRESS

Discrete Applied Mathematics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Characterization of signed line digraphs

Mukti Acharya [a], Deepa Sinha [b,*]

[a] Department of Applied Mathematics, Delhi Technological University, Main Bawana Road, Delhi-110 042, India

[b] Centre For Mathematical Sciences, Banasthali University, Banasthali - 304022, Rajasthan, India

ARTICLE INFO

Article history:
Received 2 September 2012
Accepted 4 October 2012
Available online xxxx

Keywords:
Signed graph
Signed digraph
Signed line digraph
Positive section

ABSTRACT

Given a signed digraph $S = (V(D), \mathcal{A}(D), \sigma)$ on a given digraph $D = (V, \mathcal{A})$ called the underlying digraph of S, its signed line digraph $L(S)$ is a signed digraph defined on the line digraph $L(D)$ of D by defining an arc ef in it to be negative if and only if both the arcs e and f in S are negative and oriented in the same direction through their common vertex. In this paper, we define a given signed digraph S to be a signed line digraph if there exists a signed digraph H such that $L(H) \cong S$ (read as "$L(H)$ is isomorphic to S"). We derive three structural characterizations of signed line digraphs, extending the well known characterization of line digraphs due to Hemminger (1972) [10].

1. Introduction

For standard terminology and notation in graph theory and digraph theory, not specifically defined in this paper, the reader is referred to Harary [7] and Jensen and Gutin [12], respectively. All digraphs, $D = (V(D), \mathcal{A}(D))$ with vertex set $V = V(D)$ and arc set $\mathcal{A} = \mathcal{A}(D)$ considered in this paper are finite, simple and self-loop-free.

Aigner [2] defined the 'line digraph' of a given digraph as follows. The *line-digraph* $L(D)$ of a given digraph $D = (V, \mathcal{A})$ has $\mathcal{A}$ for its vertex set and (e, f) is an arc in $L(D)$ whenever the arcs e and f in D have a vertex in common which is the *head* of e and the *tail* of f; hence, a given digraph H is called a *line digraph* if there exists a digraph D such that $L(D)$ is isomorphic to H, written $L(D) \cong H$. Harary and Norman [8] gave a characterization of line digraphs.

An ordered pair $S = (D, \sigma)$ is called a *signed digraph* with *underlying digraph* $D = (V, \mathcal{A})$, if S is obtained from D by designating each of its arcs as *positive* or *negative*.

The sets of positive arcs and negative arcs of S are denoted by $\mathcal{A}^+(S)$ and $\mathcal{A}^-(S)$, respectively. Thus, $\mathcal{A}(S) = \mathcal{A}^+(S) \cup \mathcal{A}^-(S)$. A signed digraph without orientations of its arcs (when it may be regarded as a simple undirected *signed graph*) is known as its *underlying sigraph* [3,4], denoted by S^u. It may also be regarded as a *symmetric signed digraph* in the sense that any two vertices in S are either not *linked* at all or are linked (*i.e.,* joined by an arc) by two oppositely oriented arcs of the same sign; in other words, a signed graph corresponds to the special case when its *adjacency matrix* $A(S) = (a_{ij})$ of S (*cf.:* [9]) is symmetric. A signed digraph is *all-positive* (respectively, *all-negative*) if all its arcs are positive (negative); further, it is said to be *homogeneous* if it is either all-positive or all-negative and *heterogeneous* otherwise.

Behzad and Chartrand [3] defined the *signed line graph* $L(S)$ as the signed graph in which edges of S are represented as vertices, two of these vertices are defined adjacent whenever the corresponding edges in S have a vertex in common and any such edge ef is defined to be negative whenever both e and f are negative edges in S.

* Correspondence to: Department of Mathematics, South Asian University, "Akbar Bhavan", Chanakyapuri, New Delhi-110021, India.
E-mail addresses: mukti1948@yahoo.com (M. Acharya), deepa_sinha2001@yahoo.com (D. Sinha).

0166-218X/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2012.10.005

Please cite this article in press as: M. Acharya, D. Sinha, Characterization of signed line digraphs, Discrete Applied Mathematics (2012), doi:10.1016/j.dam.2012.10.005

2. The main result

According to Hemminger and Zelinka [11], a *pseudodigraph H* is a *line digraph* if there is a pseudodigraph D such that $H \cong L(D)$. Clearly, line digraphs do not have parallel arcs; moreover, the line digraph $L(D)$ has a loop at a vertex $a \in \mathcal{A}(D)$ if and only if a is a loop in D. In this section, we shall define signed line digraphs, beginning with the following definition.

Definition 1. Given any signed digraph $S = (D, \sigma)$, its *signed line digraph* $L(S)$ is a signed digraph with $L(D)$ as its underlying digraph an arc (e, f) of which is defined negative if and only if e and f are both negative arcs in S.

It is easy to see that S is a signed line digraph of some signed digraph T if and only if $S^u \cong L(T^u)$ and $ab \in \mathcal{A}^-(S)$ if and only if $a, b \in \mathcal{A}^-(T)$. This observation motivates the next definition.

Definition 2. A given signed digraph H is a *signed line digraph* if it is isomorphic to the signed line digraph $L(S)$ of some signed digraph S.

By a *positive* (*negative*) *section* in a signed digraph $S = (D, \sigma)$, we mean a maximal edge-induced signed subdigraph consisting of only the positive (negative) arcs of S, which turns out to be simply a path (semi-path) if D is a cycle (semi-cycle) [6]. We can now give a structural characterization of signed line digraphs.

Theorem 3. *S is a signed line digraph if and only if the following conditions are satisfied:*

(i) *S^u is a line digraph.*
(ii) *For any positive arc (a, b) of S either there is no negative arc at a or there is no negative arc at b.*

Proof. *Necessity*. Suppose S is a signed line digraph, then there exists a signed digraph T such that $L(T) \cong S$ thereby implying that $S^u \cong L(T^u)$, whence (i) is seen to hold.

Now, to prove (ii), suppose (a, b) is an arc in S and there exists a negative arc (c, a) or a negative arc (a, c) at a and a negative arc (d, b) or a negative arc (b, d) at b. Then any of the following four cases may arise:

Case (i): arcs (c, a), (a, b), (b, d) exist in S;

Case (ii): arcs(a, c), (a, b), (b, d) exist in S;

Case (iii): arcs (c, a), (a, b), (d, b) exist in S;

Case (iv): arcs (a, c), (a, b), (d, b) exist in S.

Case (i). Suppose (c, a) and (b, d) are negative arcs at a and b respectively in S. This implies that c and a are adjacent negative arcs in T such that the head of the negative arc c coincides with the tail of the negative arc a in T. Also b and d are adjacent negative arcs in T such that the head of the negative arc b coincides with the tail of the negative arc d. Thus, the arc (a, b) has to be negative in S.

Case (ii). Suppose (a, c) and (b, d) are negative arcs at a and b respectively in S. This implies that a and c are adjacent negative arcs in T such that the head of the negative arc a coincides with the tail of the negative arc c in T. Also b and d are adjacent negative arcs in T such that the head of the negative arc b coincides with the tail of the negative arc d. Thus, the arc (a, b) has to be negative in S.

Case (iii). Suppose (c, a) and (d, b) are negative arcs at a and b respectively. This implies that c and a are adjacent negative arcs in T such that the head of c coincides with the tail of a in T and d and b are adjacent negative arcs in T such that the head of d coincides with the tail of b. Thus (a, b) is a negative arc in S.

Case (iv). Can be proved on similar lines.

Thus, the proof of necessary conditions follows by contraposition.

Sufficiency. Suppose (i) and (ii) hold for a given signed digraph $S = (D, \sigma)$. By (i) $D \cong L(T)$ for some digraph T.

Let $T := (V(T), \mathcal{A}(T), \sigma_T)$ be a signed digraph with $\sigma_T(a) = -$ if and only if there exists at least one negative arc at the vertex a of S, for any $a \in V(S)$. Then, an arc $(a, b) \in \mathcal{A}(L(T))$ is negative if and only if $\sigma_T(a) = -$ and $\sigma_T(b) = -$ in T such that the head of the arc a coincides with the tail of the arc b in T which is true if and only if there exist negative arcs at a and at b in S, which, in turn, is true if and only if (a, b) is a negative arc in S. This completes the proof. □

Corollary 4. *The following statements are equivalent for any signed digraph $S = (D, \sigma)$:*

(i) *S is a signed line digraph;*
(ii) *D is a line digraph and for any two vertices u, $v \in V(S)$ and for any $u - v$ semi-path in S either $u = v$ and the cycle or semi-cycle so created is not a triangle with exactly two negative arcs or the $u - v$ semi-path contains no positive section of length one unless one of its ends is u or v;*
(iii) *D is a line digraph and vertices of S can be assigned '+' or '−' such that both the ends of every negative arc receive '−' sign and the same is not true for any positive arc.*

Please cite this article in press as: M. Acharya, D. Sinha, Characterization of signed line digraphs, Discrete Applied Mathematics (2012), doi:10.1016/j.dam.2012.10.005

3. Conclusions and scope

The proof of Theorem 3 suggests that the vertices of a signed line digraph can be *marked* so that both the *ends* of every negative arc receive negative signs and no positive arc receives negative signs at both of its ends. This results into another characterization provided in Corollary 4(iii), leading to the notion of 'sign-compatibility': a signed digraph is called *sign-compatible* if its vertices can be marked such that end vertices of an arc are both negative if and only if the arc is negative.

Hence, given a signed digraph H, the set $\mathcal{L}_H = \{S : S \text{ is a signed digraph for which } L(S) \cong H\}$ may contain more than one signed digraph; the signed digraphs in $\mathcal{L}_H$ are called *signed digraph line roots* of H. As mentioned above, any signed digraph H for which $\mathcal{L}_H \neq \phi$ is sign compatible; that is, any signed line digraph is sign-compatible. However, that the converse is not true may be easily verified by the reader.

The following characterization of sign-compatible signed digraphs can be easily established.

Proposition 5. *A signed digraph S is sign-compatible if and only if its vertices can be partitioned into two subsets V_1 and V_2 such that the all-negative signed subdigraph of S is precisely the signed subdigraph induced by exactly one of the subsets V_1 and V_2.*

Next, the characterization of signed line graphs obtained by these authors [1], can be derived from Theorem 3 since every signed graph is a symmetric signed digraph.

Further, aiming towards scope, let us consider the following real-life problem. We have a goal $\mathcal{G}$ to be accomplished through completion of some p number of tasks, say, $v_1, v_2, \ldots, v_p$ in accordance with a certain *precedence relation* $\mathcal{A}$ specified *a priori* by the planning process; mathematically speaking, the consideration amounts to visualizing the ordered pair $(V, \mathcal{A})$ as a transitive digraph D where $V = \{v_1, v_2, \ldots, v_p\}$ and $(v_i, v_j) \in \mathcal{A}$ whenever the task v_i must *precede* the task v_j. As Richards [13] suggested, one might consider the "dual" notion of labelling the arcs (rather than vertices) of a digraph with tasks; following this idea, he called a digraph $D = (V, \mathcal{A})$ an *arrow diagram* of a $p \times p$ $(0, 1)$-matrix $A = (a_{ij})$ if there is a one-to-one correspondence between the set of arcs of D and the nonzero entries a_{ij} of A such that the terminal node of the arc x_i is the initial node of the arc x_j. He obtained the following characterization of square $(0, 1)$-matrices which have an arrow diagram.

Theorem 6. *A square $(0, 1)$-matrix A possesses an arrow diagram if and only if any two columns (or, equivalently, any two rows) of A are either identical or orthogonal.*

Geller and Harary [5] pointed out that Theorem 6 is in fact a characterization of a line digraph. Going further, following the same line of thought, suppose it is given that the precedence relation $\mathcal{A}$ is *dichotomically inhibited*, in the sense that each arc $(v_i, v_j) \in \mathcal{A}$ carries a 'risk factor' represented by associating the weight -1 with (v_i, v_j) if there is no risk in taking up the task v_j after having completed the task v_i and $+1$ with (v_i, v_j) if risk is perceived to exist while considering to take up the task v_j after having completed the task v_i; it is assumed, however, that the tasks have to be taken up towards accomplishing the goal $\mathcal{G}$ within the known limits of risks in each case of passing from one task to the other as per the overall plan. Is it possible to visualize salient aspects of the goal achievement plan to be able to compare it with other possible plans designed for achieving the same goal? This problem can be modelled in terms of signed digraphs. It would be interesting to investigate into the role of the study of signed line digraphs in solving this network optimization problem, wherein we anticipate that 'signed precedence relations' can be modelled as 'signed arrow diagrams', which may have a $(-1, 0, 1)$-matrix representation extending Theorem 6; specifically, we expect the truth of the following statement.

Conjecture 7. *A square $(-1, 0, 1)$-matrix A possesses a signed arrow diagram if and only if any two columns (or, equivalently, any two rows) of A are either identical or orthogonal.*

Acknowledgement

The authors are pleased to thank Dr. B.D. Acharya for sparing his valuable time for very fruitful discussions and offering useful comments throughout the preparation of this note.

References

[1] M. Acharya, D. Sinha, Characterizations of line sigraphs, Nat. Acad. Sci. Lett. 28 (1–2) (2005) 31–34. (Also, see Extended Abstract in: Electronic Notes in Discrete Mathematics, 15 (2003)).
[2] M. Aigner, On the line graph of a directed graph, Math. Z. 102 (1967) 56–61.
[3] M. Behzad, G.T. Chartrand, Line coloring of signed graphs, Elem. Math. 24 (3) (1969) 49–52.
[4] G.T. Chartrand, Graphs as Mathematical Models, Prindle, Weber and Schmidt, Inc, Boston, Massachusetts, 1977.
[5] D.P. Geller, F. Harary, Arrow diagrams are line digraphs, SIAM J. Appl. Math. 16 (6) (1968) 1141–1145.
[6] M.K. Gill, G.A. Patwardhan, A characterization of sigraphs which are switching equivalent to their line sigraphs, J. Math. Phys. Sci. 15 (6) (1981) 567–571.
[7] F. Harary, Graph Theory, Addison-Wesley Publ. Comp., Reading, Massachusetts, 1969.
[8] F. Harary, R.Z. Norman, Some properties of line digraphs, Rend. Circ. Mat. Palemo 9 (1961) 161–168.
[9] F. Harary, R.Z. Norman, D. Cartwright, Structural Models: An Introduction to the Theory of Directed graphs, Wiley, New York, 1965.
[10] R.L. Hemminger, Line digraphs, in: Proceedings of the conference Graph Theory and its Applications, Western Michigan University, Kalamazoo, Michigan, 1972, 10–13.
[11] R.L. Hemminger, Bohdan Zelinka, Line isomorphism on dipseudograph, J. Graph Theory 14 (2) (1973) 105–121.
[12] J.B. Jensen, G. Gutin, Digraphs: Theory, Algorithms and Applications, Springer-Verlag, London, 2001.
[13] P.I. Richards, Precedence constraints and arrow diagram, SIAM Rev. 9 (1967) 548–553.

Please cite this article in press as: M. Acharya, D. Sinha, Characterization of signed line digraphs, Discrete Applied Mathematics (2012), doi:10.1016/j.dam.2012.10.005

African Journal of Pharmacy and Pharmacology Vol. 6(40), pp. 2832-2836, 29 October, 2012
Available online at http://www.academicjournals.org/AJPP
DOI: 10.5897/AJPP12.682
ISSN 1996-0816 © 2012 Academic Journals

Full Length Research Paper

Clinical management of lipid profile, renal and liver function versus HbA_{1c} profile in diabetes affected patients of Vellore, Tamil Nadu, India

A. Arivarasan[1], Gaurav Rana[2], Archita Sharma[1], Manish Kumar[1], Karishma Jhang[1], Arundhati Chakraborty[1], Pravir Kumar[1,2,3] and Rashmi K Ambasta[1,2]*

[1]Functional Genomics and Cancer biology lab, Centre for Medical Engineering (CME), VIT University, Vellore, India.
[2]Department of Biotechnology, Delhi Technological University (formerly Delhi College of Engineering, Shahbad Daulatpur, Bawana Road, Delhi 110042.
[3]Neurology Department, Tufts University School of Medicine, Boston, MA USA.

Accepted 4 October, 2012

The aim of this study is to examine clinical management of lipid profile, renal and liver function versus HbA1c value correlation in diabetic population of Vellore district in South India. A total of 1300 cases from three different diabetic clinics in Vellore were taken for the study. 25% of populations were selected as control.Parameters such as age, sex, fasting blood glucose, postprandial blood glucose, lipid profile, mean plasma glucose, urine micro albumin, and HbA1c were taken into consideration. A correlation between blood glucose levels and other parameters such as lipid profile, renal and liver functionand HbA1c levels were established using meta analysis. Analysis suggests people of age group from 40 to 60 years and male population were highly affected from diabetes. Also HbA1canalysis showed only 25% of the cases had HbA1c levels between 7-7.5%, while 57% cases had HbA1c levels above 7.5%. Further, a significant positive correlation was found between postprandial blood sugar levels, total cholesterol and triglyceride levels ($p<0.0001$) which in turn indicates a chance of cardio-vascular diseases and renal impairment. The data suggest poor control of diabetes among the population and correlation study with abnormal blood glucose levels indicate complications related to diabetes including cardiovascular risk and renal impairment in patients. Hence we propose that high HbA1c levels should be considered a marker for cardiovascular risks and renal impairment and subsequent treatment must be planned in compliance to HbA1c levels.

Key words: Diabetes, HbA_{1c}, lipid profile, renal function test, liver function test, meta-analysis.

INTRODUCTION

Diabetes is a leading cause of several maladies that includes but not constrained to cardiovascular morbidity as well as mortality, and is also a significant contributor to disability in many countries (Geulayov et al., 2010). The increasing prevalence of diabetes throughout the world (Phanse et al., 2012; Zohra et al., 2012; Anbreen et al., 2012) is because of various causes viz. increase in average age of the population owing to better facilities, tremendous change in lifestyle, improved diagnostic techniques and increasing health awareness in the masses (Chou et al., 1997). The issue in diabetes control is that it largely remains under-diagnosed. In fact, a major chunk of diabetic population does not realize they are suffering from it. By and large, the average onset lag between onset and diagnosis is 7 years (Saudek et al., 2008).

The early onset of diabetes, longer suffering period out of diabetic complications and early death makes Asian population more susceptible to this disease as compared

*Corresponding author. E-mail: rashmiambasta@gmail.com. Tel: +91-9818898638.

to other regions. As a matter of fact, type II diabetes prevalence has reached to epidemic levels in Asia (Yoon et al., 2006). The number of cases of diabetes worldwide in the year 2000 among adults was estimated to be 171 million and will rise to 366 million by 2030. India is considered to be the “diabetic capital of the world”. Diabetic condition is known to be associated with environmental, behavioral, and lifestyle factors such as a sedentary lifestyle and highly rich nutrition. In addition, the daily human-environment interactions and real-life activities that cause an individual's blood glucose to fluctuate remain relatively unexplored owing in part to data collection challenges (Doherty, 2011). Global prevalence of diabetes was about 2.8% in 2000 and is estimated to be around 4.4% by 2030. It has been estimated that in the 30-year period, the prevalence of diabetes in India would increase by 195%, which is the highest in the world. The International Diabetes Federation estimates the total number of diabetic subjects to be around 40.9 million in India and this is further set to rise to 69.9 million by the year 2025 (Murugesh and Mani, 2011).

Public health and managed-care organizations typically use multiple data sources to track and monitor chronic diseases such as diabetes (Saydah et al., 2004). Several efficient but underutilized preventive-care practices have proven to reduce complications related to diabetes. This has been largely established with the support of several surveillance projects. However, strategies to identify individuals with diagnosed diabetes from various data sources have not been systematically evaluated. The focus of care for people with diabetes has shifted from hospital to general practice. Many practices now offer diabetes care via dedicated mini-clinics, shared care schemes or opportunistically (Pierce et al., 2000).

The main objective of our survey was to gather data of blood glucose level as well as lipid profile (Giansanti et al., 1999; Smith, 2007; Gotto, 2007; Khan et al., 2007; Selvin et al., 2006) from which we drew results such as: which areas are getting good, bad or average diabetes health care. This survey reconsiders the criteria for diagnosing diabetes and discusses screening criteria to make case finding easier for clinicians as well as patients. A cross sectional study of people belonging to all ages with diabetes was performed in this survey and varying results were so obtained. This survey shall present a clear detail about the prevalence of diabetes in Vellore district of Tamil Nadu in South India.

MATERIALS AND METHODS

Data were collected from three different diabetic clinics in Vellore. In total, 1300 cases were taken into consideration out of which 25% were taken as control population. Parameters such as age, sex, fasting blood glucose, postprandial blood glucose, lipid profile (total cholesterol, blood serum triglyceride (TGL), low-density lipoprotein (LDL), high-density lipoprotein (HDL), very low-density lipoprotein (VLDL), Cho/HDL, LDL/HDL), liver function tests (total protein, albumin, globulin A/G, total bilirubin), renal function test (plasma urea, creatinine, serum glutamic oxaloacetic transaminase (SGOT), serum glutamic pyruvic transaminase (SGPT), sodium (Na^+), potassium (K^+), chloride (Cl^-) and bicarbonate ions value), and HbA_{1c} were taken into consideration. The data was analyzed on the basis of age groups and sex in which individuals starting from age 10 were taken into consideration and their control over diabetes was analyzed on the basis of HbA_{1c} test values.

Data was analyzed using Chi-square ($\chi 2$) test. Correlation was also established between blood glucose levels and other parameters such as lipid profile, mean plasma glucose, urine micro-albumin and HbA_{1c} levels. To better understand the correlation between age, sex and HbA_{1c} levels of the diabetic population, certain techniques were employed. The parameters were also correlated were the lipid profile, urine micro-albumin and mean plasma glucose with impaired fasting and postprandial blood glucose levels. These were extensively recorded in control and diabetic patient groups. The clinics which aided the survey by supplementing the diabetic data are: Vellore Diabetic Clinic # 1, Vellore, India (Data size- 600); Vellore Diabetic Clinic # 2, Vellore, India (Data size- 400); Vellore Diabetic Clinic # 3, India (Data size- 300). All 1300 cases were diabetic but 25% were normalized after metformin treatment. Therefore, they were excluded from the study. The duration of the diabetic disease was 4 - 5 years and the patients were on metformin. Any major complications were not reported in the small duration.

Meta-analysis and statistical tests

Data analysis was conducted with statistical package ‘GraphPad Prism 5.0’, Microsoft Excel and meta-analysis tool. The χ^2 test, ANOVA test and Pearson coefficient were used to measure the association among the different variables. The results were generated in the form of averages, standard deviation and their 95% confidence interval (CI). Correlation was established using Pearson’s coefficient between fasting and postprandial blood glucose levels and various parameters such as total cholesterol, triglyceride levels, LDL, VLDL, urine micro-albumin, and mean plasma glucose. Type of correlation was established using coefficient of determination, r^2. The p values of < 0.05 were defined as significant.

RESULTS

Diabetes occurrence is dependent on age and sex

Out of the total data size of 1300, 557 were male diabetic patients and 402 were female diabetic patients (Table 1). The 341 individuals were close to normal, hence they were excluded from the study. The data was divided according to the age groups and sex. The tabulated data showed a peak in the age group of 50 - 60 years. The Figure 1 obtained showed that the population of male diabetes was more than the female diabetic in Vellore and also showed a peak in the age group of 50 - 60. Thus, it can be safely inferred that male individuals are more prone to diabetes in the middle age group of 40 - 60 years (Table 2). Even females suffering from diabetes were observed to be higher in age group of 40 - 60 years, but still the occurrence rate was found to be lower than the males. The difference in occurrence of diabetes between males and females can be due to difference in

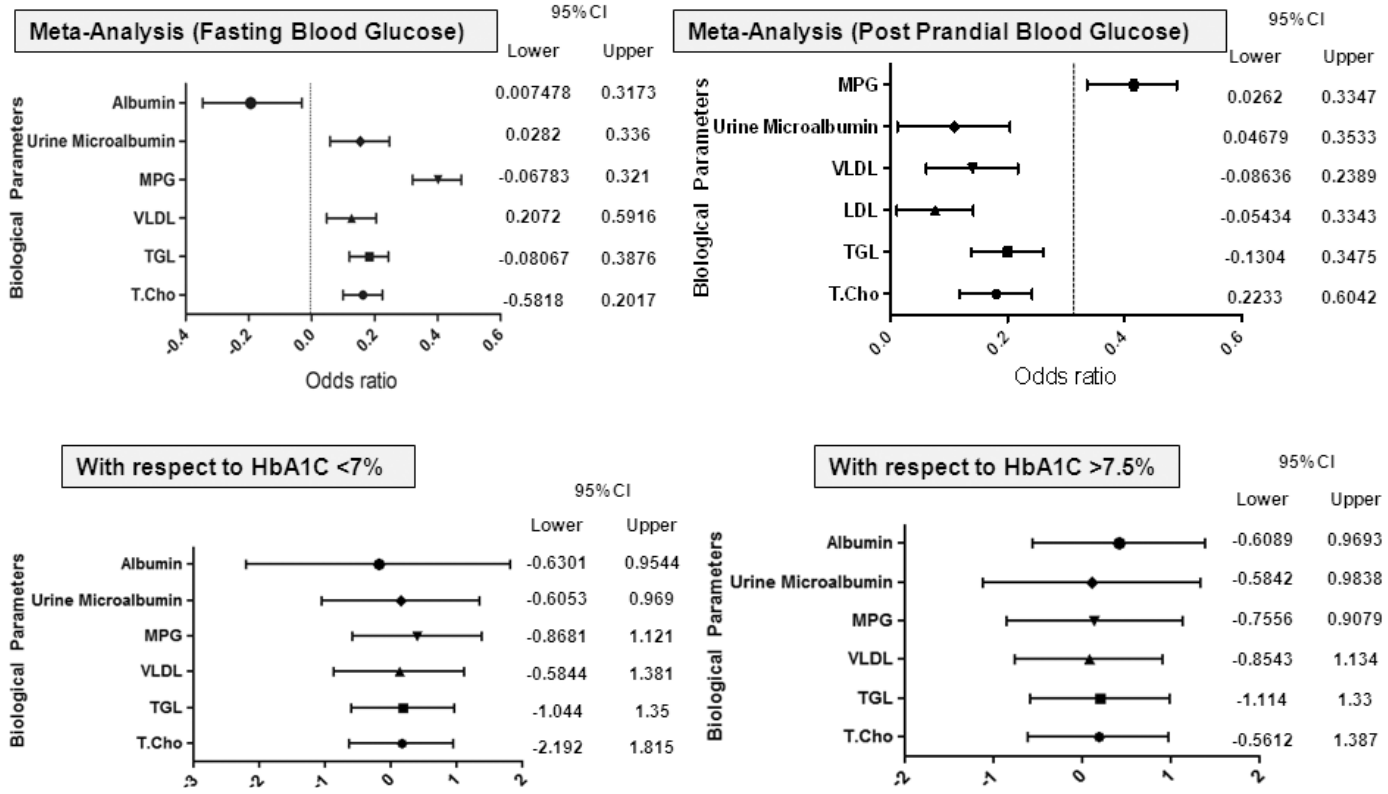


Figure 1. Meta-analysis of various parameters compared in patient's fasting glucose levels.

Table 1. Data on age- and sex-dependent diabetic patients

Age	Diabetic male	Diabetic female
10 - 20	0	3
20 - 30	17	9
30 - 40	82	66
40 - 50	169	126
50 - 60	190	122
60 - 70	74	66
70 - 80	24	10
80 - 90	1	0

Table 2. Results on the sex-dependent HbA_{1c} value fluctuation.

Parameter	HbA_{1c} values		
	<7%	7-7.5%	>7.5%
Diabetic male	96	101	226
Diabetic female	68	66	60

their lifestyle. The survey findings suggest that the males must do more exercise and reduce sedentary lifestyle so as to come over the bias of high susceptibility to diabetes. Stress could also be one of the prominent factors for higher occurrence of diabetes in males.

HbA_{1c} value in diabetic patients

The HbA_{1c} values were categorized based on the findings that the HbA_{1c} level for a diabetic patient should be around 7-7.5 %. The data was grouped under three categories viz. <7%, 7-7.5 and >7.5% HbA_{1c} levels. Amongst these, the category comprising of >7.5% showed the maximum especially for diabetic male population. Therefore, it can be safely concluded that the majority of the diabetic population in Vellore region lacks proper control over the glucose levels. Furthermore, the health care facilities provided for the control of diabetes might not be adequate.

Correlation between blood glucose and renal/liver function and lipid profile in diabetic patients

Statistical studies were conducted to correlate fasting blood glucose to total cholesterol and lipid profile in the populations. In addition, correlation was also established between postprandial glucose and total cholesterol con-

Table 3. Results on the correlation between fasting and postprandial blood glucose levels with total cholesterol, triglyceride (TGL) levels, low density lipoprotein (LDL), very low density lipoprotein (VLDL), mean plasma glucose (MPG) and urine micro-albumin.

Parameter	Pearson (r)	95% CI	p value	R squared
Fasting glucose and Total cholesterol	0.1628	0.09980 to 0.2245	< 0.0001	0.02651
Postprandial glucose and total cholesterol	0.1810	0.1181 to 0.2423	< 0.0001	0.03274
Fasting glucose and TGL	0.1826	0.1199 to 0.2438	< 0.0001	0.03333
Postprandial and TGL	0.2006	0.1381 to 0.2615	< 0.0001	0.04024
Postprandial and LDL	0.07647	0.01068 to 0.1416	0.0228	0.005848
Fasting glucose and VLDL	0.1271	0.04807 to 0.2046	0.0017	0.01616
Postprandial and VLDL	0.1406	0.06145 to 0.2179	0.0005	0.01976
Fasting glucose and MPG	0.4013	0.3211 to 0.4758	< 0.0001	0.1610
Postprandial glucose and MPG	0.4157	0.3361 to 0.4894	< 0.0001	0.1728

centration as tabulated in Table 3. The same table also depicts correlation between fasting and postprandial glucose and triglycerides level in diabetic patients. However, no significant correlation was observed with other noted parameters; hence they were excluded from the analysis.

DISCUSSION

Several studies have analyzed and reported the clinical management of diabetes at zonal level, however, the Vellore region has largely remained ignored in this aspect (Annette, 2012; Evelyn et al., 2012; Health Educ. Res., 2012). From the values obtained in Table 3, we were able to find a positive correlation between various parameters. There was a significant and positive correlation between fasting blood glucose level and total cholesterol (r =0.1628, 95%CI= 0.09980 to 0.2245, p= <0.0001) and postprandial blood glucose and total cholesterol (r = 0.1810, 95% CI= 0.1181- 0.2.423, p= <0.0001). Also, there was a positive correlation between fasting blood glucose and triglyceride (TGL) levels (r = 0.1826, 95%CI = 0.1199 to 0.2438, p= <0.0001); and postprandial blood glucose and TGL levels (r = 0.2006, 95%CI= 0.1381 to 0.2615, p= <0.0001). Positive correlation was also found between postprandial LDL levels (r = 0.07647, 95% CI= 0.01068 to 0.1416, p = 0.0228) and fasting blood glucose and VLDL (r = 0.1271, 95%CI= 0.04807 to 0.2046, p= 0.0017), and postprandial blood glucose and VLDL (r = 0.1406, 95% CI= 0.06145 to 0.2179, p=0.0005).

The positive correlation between fasting blood glucose and postprandial blood glucose with total cholesterol, triglyceride levels, LDL and VLDL indicate the risk of cardiovascular diseases in the patients. Thus, we can conclude that the diabetic populations with higher blood glucose levels in Vellore are more prone to cardio-vascular diseases. The correlation was also established between fasting blood glucose and mean plasma glucose (MPG) (r = 0.4013, 95%CI= 0.3211 to 0.4758, p<0.0001) and postprandial blood sugar and MPG (r = 0.4157, 95% CI= 0.3361 to 0.4894, p= <0.0001). Thus, impaired blood glucose levels were found to be responsible for high MPG levels. We also found a correlation between fasting blood glucose and urine micro-albumin (r = 0.1544, 95% CI= 0.5872 to 0.2472, p= 0.0017) and postprandial blood glucose and urine micro-albumin (r = 0.1093, 95% CI= 0.1203 to 0.2044, p= 0.0277). This correlation was also found to be positive and significant and indicated the risk of renal disorders due to impaired blood glucose levels in the diabetic population in Vellore. However, no significant correlation was found for liver dysfunction.

Herein, we concluded that random sampling of diabetic population reflects that propensity of diabetes is much higher in male (between 40 - 60 years age group) than the female population in Vellore. Furthermore, we observed that the percentage of glycosylated blood (HbA1c) value in mass population is beyond the normal range of 6.5 - 7%, indicating a sedentary life style, more calorie and glucose intake as well as poor awareness towards the health condition (Colberg, 2012; Cinar et al., 2012; Grandy et al., 2012; Tsang, 2012) that could be attributed to financial constraints. In addition, correlation between diabetes and cardiovascular risk is well established. In our study, we found that the lipid profile is altered in diabetic patients than the normal population. Also, the renal failure and diabetic nephropathy is a lethal downstream event in the diabetic population that is clearly observed with high urine micro-albumin levels in Vellore population. Therefore, these findings are beneficial to treat the diabetic patients more effectively and can help to combat diabetic-related disorders since the management of diabetes is poor in Vellore, Tamil Nadu, India. The diabetic individuals may further be suggested for self-diabetes management programs. The government bodies and concerned authorities may launch training and educational programs to raise awareness in the masses (Lynch et al., 2012).

Conclusion

We have found that there is correlation between higher HbA1C value and cardiovascular risk, renal function and

liver function test. Therefore, this region needs to focus more on diabetes management and diabetes-related disorders.

ACKNOWLEDGEMENTS

We would like to thank the Vellore hospitals, who have contributed human diabetic data for this study. The authors also thank the management of VIT for their continuous support.

REFERENCES

Anbreen A, Muhammad Y, Muhammad Z, Tanveer HB, Ameer FZ, Zafar IK, Shazia N, Kulsoom GA, Bushra P, Kafeel A, Muhammad KM, Saira H, Sajjad A, Muhammad UT, Ghulam H (2012). A comparative study on the status of Zn and Cu in diabetic and non diabetic males in Punjab, Pakistan. Afr. J. Pharm. Pharmacol., 6(20): 1482-1486.

Annette LM (2012). Changes and Consistencies in Diabetes Education Over 5 Years: Results of the 2010 National Diabetes Education Practice Survey. The Diabetes Educator, 38: 35-46,

Chou P, Li CL, Kuo HS, Hsiao KJ, Tsai ST (1997). Comparison of the prevalence in two diabetes surveys in Pu-Li, Taiwan, 1987-1988 and 1991-1992. Diabetes Res. Clin. Pract., 38(1): 61-67.

Cinar AB, Oktay I, Schou L (2012). Relationship between oral health, diabetes management and sleep apnea. Clin. Oral Investig, May 31.

Colberg SR (2012) Physical activity: the forgotten tool for type 2 diabetes management. Front Endocrinol. (Lausanne), 3:70. Epub 2012.

Doherty ST (2011). Exploring Blood Glucose Variation over Geographical Space Diabetes Technol Ther. 2012;14(3):276-84.

Evelyn YH, Catherine A. Chesla, Kevin MC (2012). Health Communication With Chinese Americans About Type 2 Diabetes. The Diabetes Educator, 38: 67-76,

Geulayov G, Goral A, Muhsen K, Lipsitz J, Gross R (2010). Physical inactivity among adults with diabetes mellitus and depressive symptoms: results from two independent national health surveys. Gen Hosp Psychiatry. 2010 Nov-Dec; 32(6): 570-576.

Giansanti R, Rabini RA, Romagnoli F, Fumelli D, Sorichetti P, Boemi M, Fumelli P (1999) Coronary heart disease, type 2 diabetes mellitus and cardiovascular disease risk factors: a study on a middle-aged and elderly population. Arch. Genontol. Geriatr., 29: 175–182.

Gotto AM Jr (2007). Cardiologist's role in improving glucose control and global cardiovascular risk in patients with type 2 diabetes mellitus. Am. J. Cardiol., 99: 3–5.

Grandy S, Fox KM, Bazata DD (2012). Association of Self-Reported Weight Change and Quality of Life, and Exercise and Weight Management Behaviors Among Adults with Type 2 Diabetes Mellitus: The SHIELD Study; Cardiol Res Pract.;2012:892564. 7p.

Khan HA, Sobki SH, Khan SA (2007). Association between glycemic control and serum lipids profile in type 2 diabetic patients: HbA1c predicts dyslipidemia. Clin. Exp. Med., 7: 24-29.

Lynch EB, Fernandez A, Lighthouse N, Mendenhall E, Jacobs E. Concepts of diabetes self-management in Mexican American and African American low-income patients with diabetes. Health Educ Res. 2012; 27(5):814-24.

Murugesh S, Mani D (2011). Prevalence of diabetes in Vellore district, Tamil Nadu, India. J. Chem. Pharm. Res., 3(4): 684-696.

Phanse M, Padhye S, Patil M, Takawale A, Navghare V (2012). Effect of thespesone-vanadium complex in alloxan induced diabetic rats. Afr. J. Pharm. Pharmacol., 6(10): 692–697.

Pierce M, Agarwal G, Ridout D (2000). A survey of diabetes care in general practice in England and Wales. Br. J. Gen. Pract, 50(456): 542-545.

Saudek CD, Herman WH, Sacks DB, Bergenstal RM, Edelman D, Davidson MB (2008). A new look at screening and diagnosing diabetes mellitus. J. Clin. Endocrinol. Metab., (7): 2447-53. Epub 2008 May 6. Rev.

Saydah SH, Geiss LS, Tierney E, Benjamin SM, Engelgau M, Brancati F (2004). Review of the performance of methods to identify diabetes cases among vital statistics, administrative, and survey data. Ann. Epidemiol., 14(7):507-16.

Selvin E, Wattanakit K, Steffes MW, Coresh J, Sharrett AR. (2006). HbA1c and peripheral arterial disease in diabetes: the Atherosclerosis Risk in Communities study. Diabetes Care, 29: 877-882.

Smith SC Jr (2007). Multiple risk factors for cardiovascular disease and diabetes mellitus. Am. J. Med., 120:S3–S11

Tsang MW (2012). The management of type 2 diabetic patients with hypoglycaemic agents. ISRN Endocrinol., 2012:478120.

Yoon KH, Lee JH, Kim JW, Cho JH, Choi YH, Ko SH, Zimmet P, Son HY (2006). Epidemic obesity and type 2 diabetes in Asia. Lancet, 368(9548): 1681-8. Rev.

Zohra G, Khaled H, Mongi S, Zouheir S, Khaled MZ, Abdelfattah EF, Ahmed H (2012). Effect of Nigella sativa seeds on reproductive system of male diabetic rats. Afr. J. Pharm. Pharmacol., 6(20): 1444-1450.

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Energy Procedia 14 (2012) 695 – 701

Energy Procedia

2nd International Conference on Advances in Energy Engineering (ICAEE 2011)

Comparison of Various Auxiliary Signals for Damping Subsynchronous Oscillations Using TCR-FC

Sanjiv Kumar[a*], Dr. Narendra Kumar[b] ,Vipin Jain[c]

[b]Department of Electrical Engineering,Delhi Technological University, Delhi-110042 India.
[a,c]Research Scholar, Delhi University, Delhi-110042 India.

Abstract

Because of large-capacity, long-distance and cross-region power transmission in our country's power grid, the subsynchronous oscillation (SSO) will arise from the use of series capacitor compensation and HVDC. Damping subsynchronous oscillation using static VAR compensators (SVC) is investigated in this paper.TCR-FC (Thyristor Controlled Rectifier with Fixed Capacitor) is a well known combination to improve voltage stability. Supplementary signals such as variation in reactive power, variation in frequency, variation in active power, variation in current can be used to enhance the dynamic response of the system. Their derivatives also can be used for better performance. In this paper three signals are compared and it is shown that Deviation in Reactive power gives best performance.

Keywords- FACTS; TCR-FC; Reactive power deviation; active power deviation; frequency deviation

1. Introduction

Control changing the network parameters is an effective method of improving transient stability. Flexible ac transmission system (FACTS) controllers due to their rapid response are suitable for transient stability control since they can bring about quick changes in the network parameters. Transient stability control involves changing the control variables such that the system state enters the stability region after a large disturbance [1].

Control by changing the network parameters is an effective method of improving transient stability. Flexible ac transmission system (FACTS) controllers due to their rapid response are suitable for transient stability control since they can bring about quick changes in the network parameters. Transient stability control involves changing the control variables such that the system state enters the stability region after a large disturbance.

* Corresponding author. Tel.: +91-9410436382
E-mail address: activesanjiv007@rediffmail.com

doi:10.1016/j.egypro.2011.12.887

TCR-FC (Thyristor Controlled Rectifier with Fixed Capacitor) is a well known combination to improve voltage stability. Supplementary signals are used to improve dynamics of power system, i.e to reduce power oscillations etc. Damping of power system oscillation plays an important role not only in increasing the transmission capability but also for stabilization of power system conditions after critical faults, particularly in weakly coupled networks. Series compensation has been widely used to enhance the power transfer capability. However, series compensation gives rise to dynamic instability and sub synchronous resonance (SSR) problems. Many preventive measures to cope with this dynamic instability problem in series compensated lines have been reported in literature.

These supplementary signals may be deviation in Reactive power, deviation in frequency, deviation in bus angle voltage, deviation in active power.etc. [3] [4]. Damping of power system oscillations plays an important role not only in increasing the power transmission capability but also for stabilization of power system conditions after critical faults [12]. In this paper deviation in reactive power, active power & frequency is used as Supplementary signal.

Subsynchronous resonance is addressed in three categories (i) induction generator effect (ii) torsional effect (iii) torque amplification. In all cases SSR is due to the interaction of a series capacitor with turbine generator. The first two types are caused by a steady state disturbance, while the third is excited by transient disturbance. Flexible AC transmission system (FACTS) technology provides unprecedented way for controlling transmission grids and increasing transmission capacity [7–9]. FACTS controllers have the flexibility of controlling both real and reactive power which could provide an excellent capability for improving power system dynamics. Several studies have investigated the potential of using this capability in mitigating SSR of series capacitive compensated transmission grids [10–15].

Two IEEE benchmark models have been proposed by the IEEE-SSR Working Group. These benchmark models have obtained world-wide acceptance and are extensively used for the study of different proposed damping devices SSR countermeasures [16-17].

The use of the thyristor controlled series capacitor (TCSC), static synchronous compensator and static synchronous series compensator in their balanced mode of operations has been implemented and/or studied as means for damping SSR. Generally FACTS controllers are used for power flow control and voltage stability [18-20]. Very less research has been carried out for damping of SSR using FACTS devices. In our research paper we have shown that UPFC is an effective FACTS device for damping of SSR. [21-28].

In this research paper we have developed a transmission system in MATLAB very similar to the IEEE first benchmark model.

It is helpful to look at the two SSR types within the classification which results from different sets of assumptions in our simplified model:

a) *Constant current field winding:* Only torsional interaction is present as a result of which currents and voltages at subsynchronous frequency in the stator and the shaft torque grow.

b) Damper winding on the rotor with zero initial current:

Only induction generator effect is present as a result of which currents and voltages at subsynchronous frequency in the stator, current in the rotor and the shaft torque grow.

c) Constant current field winding, constant synchronous speed: Interaction at subsynchronous frequency between the stator and rotor circuits stops. If there is no resistance in the stator, the subsynchronous currents and voltages resulting from initial conditions continue to exist without growing.

Also the electromagnetic torque component on the generator rotor is present, but since with our constant speed assumption we have actually placed an independent torque source on the generator rotor which completely compensates this torque, the mechanical oscillations do not grow.

d) Damper winding with zero initial current, constant synchronous speed: Interaction at synchronous frequency between the stator and rotor circuits stops. The electrical induction effect on the stator side results in currents and voltages at subsynchronous frequency to grow.

2. Study System

The study system consists of a steam turbine driven synchronous generator supplying bulk power to an infinite bus over a long transmission line. The study system, shown in Fig.1 consists of one synchronous generator, two transformers T_1 & T_2, one series

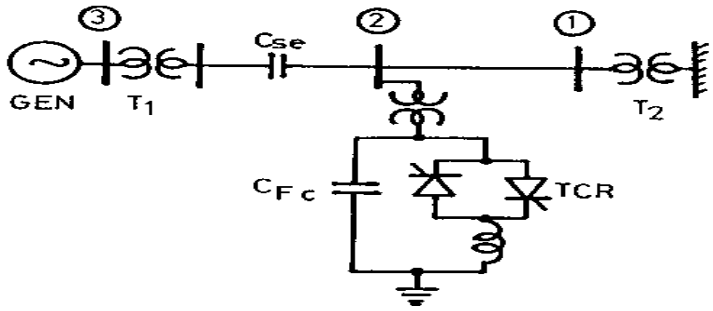


Fig.1

capacitor and a TCR-FC in the middle of line. Modelling of above system & initial conditions are given in [1]. Block diagram of TCR-FC is shown in Fig.2. This block diagram includes firing control system and represented as a first order model having gain K and time constants T_1 and T_2.

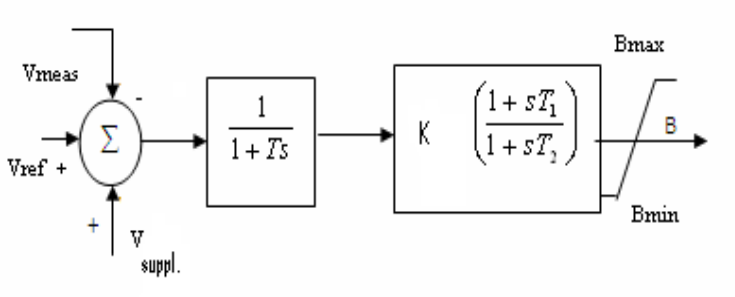


Fig.2

The controller send firing control signals to the thyristor switching unit to modify the equivalent susceptance of the TCR. In Fig. 2 V_{meas} is bus voltage at TCR –FC bus. V_{suppl} is change in voltage due to auxiliary signal deviation (reactive power, active power or frequency deviation) at TCR –FC bus.

3. Fault Simulation

In synchronous machine initial power is 2 pu. Suddenly it increased to 2.5 pu.

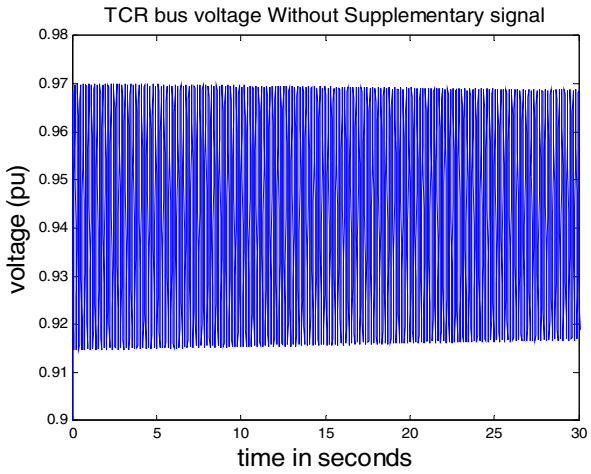


Fig.4

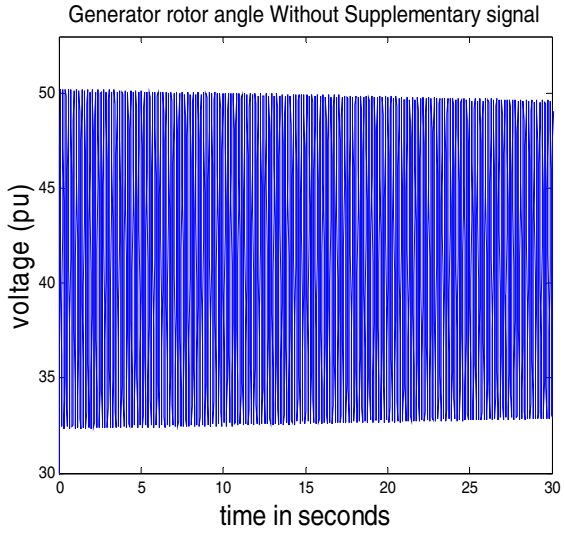


Fig. 5

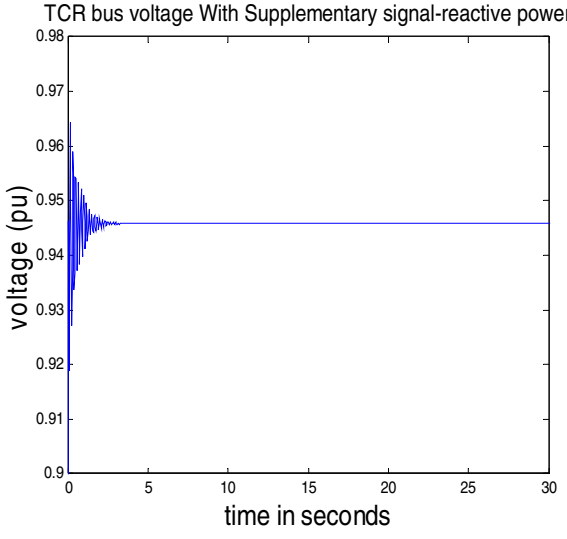


Fig. 6

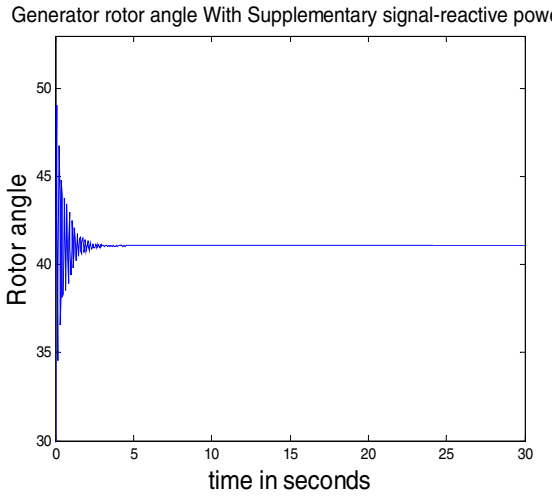


Fig. 7

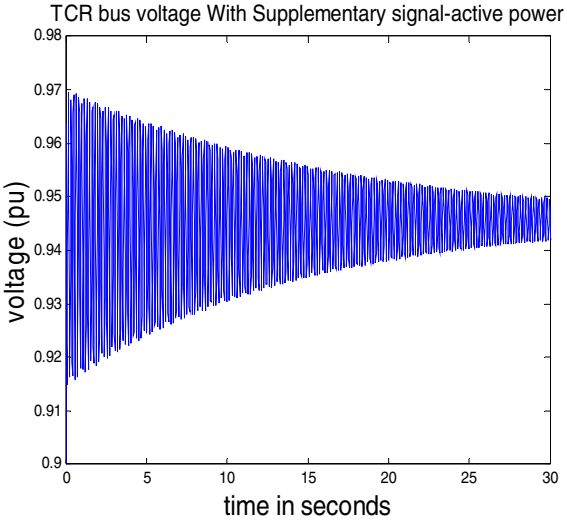


Fig. 8

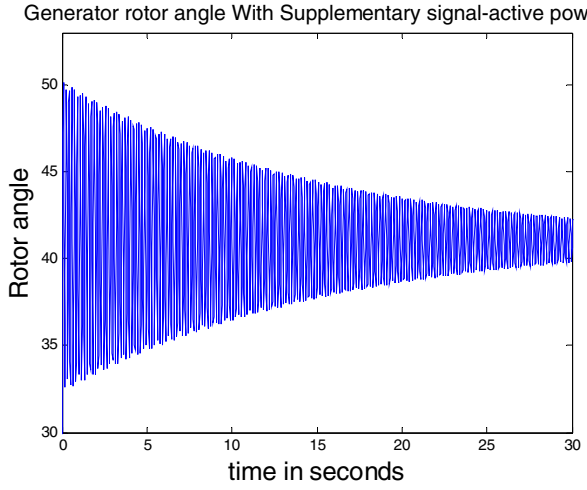


Fig. 9

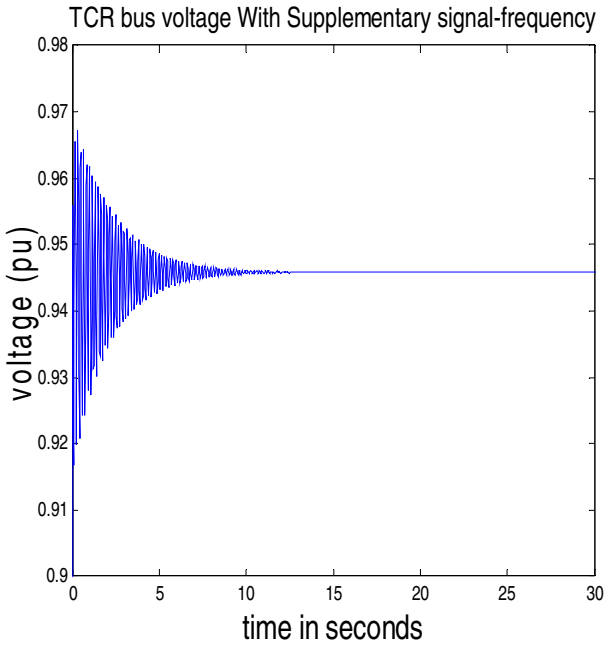


Fig. 10

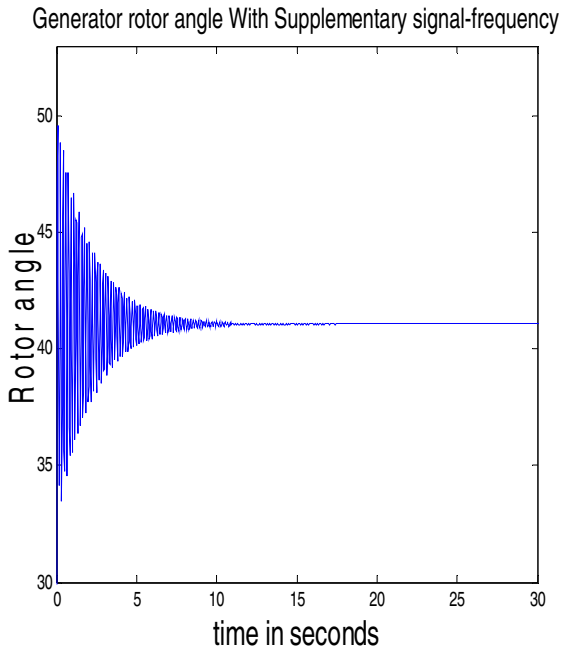


Fig.11

4. Results

(a) Without any supplementary controller - Voltage stability - Fig.4 Rotor stability - Fig.5
(b) With supplementary controller (auxiliary signal-reactive power) Voltage stability-Fig.6,Rotor stability Fig.7
(c) With supplementary controller (auxiliary signal-active power) Voltage stability -Fig. 8, Rotor stability Fig.9
(d) With supplementary controller (auxiliary signal-frequency deviation) Voltage stability -Fig. 10, Rotor stability Fig.11

5. Conclusion

In this paper the effectiveness of combined voltage and reactive power with TCR-FC auxiliary controllers have been evaluated for damping the subsynchronous oscillations in a given series compensated power system.Supplementary signal deviation in reactive power (Fig.6 & 7) gives best result, then frequency deviation gives little better results and out of three signals deviation in active power gives least result. No doubt that without Supplementary signal (Fig. 4 & 5) results are poorest. The response curves of terminal voltage, TCR-FC bus voltage, generator torque angle show a remarkable improvement and their oscillations die down effectively.The further work is going on with combining two signals in one controller.

Acknowledgements

The work presented in this paper has been performed under the AICTE R&D Project, “Enhancing the power system performance using FACTS devices” in the Flexible AC Transmission Research Laboratory at Delhi Technological University formerly Delhi College of Engineering, Delhi (India).

References

[1] N.Kumar, S.T.Nagarjan, "A SVS control strategy for damping Torsional Oscillation due to SSR in a series compensated system" Institution of Engineers (india), Vol. 91. March 2011
[2] Anderson,P.M,Agarwal,B.L.,Van Ness,J.E"Subsynchronous Resonance in Power System', IEEE Press.
[3] P.Kundur, Power System Stability and Control, Mc Graw Hill, 1994.
[4] Y. Wang, R.R. Mohler "Variable structure facts controller for power system transient stability" Transactions on Power Systems, Vol. 7, No. 1, February 1992
[5] W.Sae-Kok,A.Yokoyama,S.C.Verma " Excitation Control System Design of Rotary Type Frequency Converter for Performance Improvement of Power System Dynamics" IEEE Trans. on energy conversion, vol. 21, no. 1, March 2006
[6] M. Noroozian, L. Angquist, M. Ghandhari, G. Anderson" Improving power system dynamics by series-connected facts devices" IEEE Transactions on Power Delivery, Vol. 12, No. 4, October 1997.
[7] N.Karpagam, D.Devaraj "Application of GA for SVC –FACTS Controller for Power System Transient Stability Improvement" International Journal of Electrical Power and Energy Systems Engineering 2:2 2009
[8] Harbans Nakra, R. L.ewh Vaughan, Charles Gagnon-" Real-Time Simulator for Power System Dynamics Studies" IEEE Transactions on Power Systems, Vol. 10, No. 2, May 1995
[9] D. J. Trudnowski, M. K. Donnelly, J. F. Hauer –"Estimating Damping Effectiveness of BPA's Thyristor Controlled Series Capacitor by Applying Time and Frequency Domain Methods to Measured Response" IEEE Transactions on Power Systems, Vol. 11. No. 2, May 1996
[10] M. L. Crow , J. G. Chen"The multirate simulation of facts devices in power system dynamics" IEEE Transactions on Power Systems, Vol. 11, No. 1, February 1996
[11] Varma, R. K.; Auddy, S.; Semsedini, Y, "Mitigation of Subsynchronous Resonance in a Series-Compensated Wind Farm Using FACTS Controllers" IEEE Transaction on Power Delivery Page(s): 1645-1654 Vol.3,2008
[12] S.K. Gupta, Narendra Kumar, "Controlled Series Compensation in Coordination with Double Order SVS Auxiliary Controller and Induction Machine for repressing the Torsional Oscillations in Power system". Electric Power System Research 62, 93-103

Appendix A.

Generator data: 1110MVA, 22kV, R_a= 0.0036, X_L = 0.21
T_{do}'=6.66, T_{qo}'=0.44, T_{do}''=0.032, T_{qo}''=0.057s
X_d = 1.933, X_q = 1.743, X_d' =0.467, X_q' = 1.144, X_d'' = 0.312, X_q'' = 0.312 p.u.
IEEE type 1 excitation system:
T_R=0, T_A=0.02, T_E=1.0, T_F=1.0s, K_A=400, K_E=1.0; K_F=0.06 p.u.
V_{fmax}=3.9, V_{fmin}=0, V_{rmax}=7.3, V_{rmin}= -7.3
Transformer data:
R_T=0, X_T=0.15 p.u. (generator base)

Transmission line data:

Voltage 400kV, Length 600km, Resistance R=0.034Ω / km, Reactance X=0.325 Ω / km
Susceptance B_c=3.7μ mho / km

SVS data:

Six-pulse operation:

T_M=2.4, T_S=5, T_D = 1.667ms, K_1= 1200, K_P = 0.5, K_D = 0.01

Torsional spring-mass system data

Mass	shaft	Inertia H (s)	Spring constant K (p.u. torque/rad)

HP		0.1033586	
	HP-IP		25.772
IP		0.1731106	
	IP-LPA		46.635
LPA		0.9553691	
	LPA-LPB		69.478
LPB		0.9837909	
	LPB-GEN		94.605
GEN		0.9663006	
	GEN-EXC		3.768
EXC		0.0380697	

J Mater Sci: Mater Electron
DOI 10.1007/s10854-012-0963-7

Effect of sintering temperature on the grain growth and electrical properties of barium zirconate titanate ferroelectric ceramics

Priyanka A. Jha · A. K. Jha

Received: 30 August 2012 / Accepted: 29 October 2012

Abstract In the present work, structural, dielectric and ferroelectric properties of barium zirconate titanate ferroelectric ceramics have been investigated. The specimens were synthesized using a solid state reaction technique. The XRD analysis reveals that the synthesized compound was formed with no secondary phases. As the sintering temperature increases from 1,200 to 1,300 °C, the average grain size is observed to increase from ~0.39 to ~6.15 μm. The dielectric measurements as a function of temperature show a decrease in Curie temperature (T_C) on increasing the sintering temperature. The decrease in Curie temperature is attributed to the substitution of Zr^{+4} whose ionic radius is larger than Ti^{+4}. A large increase in the dielectric constant with the increase in grain size is observed. The remanent polarization is also observed to increase with the increase in grain size.

1 Introduction

Barium titanate is a very useful material owing to its high dielectric constant and lead free composition [1, 2]. This material has been used for a wide range of scientific and industrial applications such as capacitors, ultrasonic transducers, pyroelectric infrared sensors, etc. At the Curie temperature (T_c) there is a sharp phase transition and its structure undergoes a phase change from tetragonal to cubic. The phase transition temperature can be modified by the partial substitution of either Ba-ions or Ti-ions (A-site substitution or B-site substitution). A-site substitution with the same valence increases or decreases the Curie temperature without any significant broadening of the transition [3]. With the B-site substitution, the ferroelectric domains, which are associated with a cooperative off—center displacement of Ti^{+4} ions in their TiO_6 octahedra, are disrupted, which often leads to a broadening of the transition around the T_c [4].

Recently, barium zirconate titanate (BZT) has attracted considerable attention due to its applications as piezoelectric transducers, DRAM and tunable microwave devices including tunable filters, phase shifters, antennas, etc. [5–11]. The nature of phase transition temperature is known to change strongly with Zr content [12]. A detailed literature survey reveals that there is a disagreement about the extent of the effect of Zr^{+4} substitutions in $BaTiO_3$ on the Curie temperature [13–18]. This prompted the authors to study this compound and investigate the effect of sintering temperature on the phase formation and the nature of the phase transition in $BaZr_{0.05}Ti_{0.95}O_3$.

2 Experimental

Polycrystalline samples of the composition $BaZr_{0.05}Ti_{0.95}O_3$ were prepared by the solid state reaction method by taking high purity $BaCO_3$, TiO_2 and ZrO_2 (all from M/s Aldrich, USA) in their stoichiometric proportions. The

Electronic supplementary material The online version of this article (doi:10.1007/s10854-012-0963-7) contains supplementary material, which is available to authorized users.

P. A. Jha · A. K. Jha (✉)
Thin Film and Material Science Laboratory, Department of Applied Physics, Delhi Technological University (Formerly Delhi College of Engineering), Bawana Road, Delhi 110042, India
e-mail: prof.akjha@gmail.com

A. K. Jha
Department of Applied Sciences, A.I.A.C.T.R., Geeta Colony, Delhi 110031, India

Published online: 08 November 2012

powder mixtures were thoroughly ground in an agate mortar and passed through a sieve of 60 μm mesh. The powder mixture was then calcined at 1150 °C for 2 h in an alumina crucible. The calcined mixtures were ground and admixed with polyvinyl alcohol (PVA) as a binder and then pressed at 150 MPa into disk shaped pellets. The pellets were then sintered at 1,200, 1,250 and 1,300 °C for 2 h in air. X-ray diffractograms of the sintered samples were recorded using a X-ray diffractometer (D8 Advance, Bruker, Karlsruhe, Germany) with CuKα ($\lambda = 1.5406$ A°) radiations in the range $10° \leq 2\theta \leq 70°$ at a scanning rate of 2°/minute. The microstructural studies of the samples were carried out using Scanning Electron Microscope (Hitachi S-3700N, Tokyo, Japan) after gold coating. The grain size has been calculated using Heyn's line Intercept Method as defined by ASTM E112-96 with the help of ImageJ software. The sintered pellets were polished to a thickness of 1 mm and coated with silver paste on both sides to act as electrodes and then finally cured at 300 °C for 30 min. The dielectric measurements were carried out in the temperature range of room temperature to 300 °C at 1, 10 and 100 kHz using a high precision LCR meter (Agilent 4284A, Santa Clara, CA) at oscillation amplitude of 1 volt. Polarization (P) and Electric Field (E) hysteresis loops were recorded at room temperature using an automatic P–E loop tracer based on Sawyer–Tower circuit.

3 Results and discussions

3.1 XRD analysis

X-ray diffractograms of the studied specimens are shown in Fig. 1. Lattice parameters were calculated from the diffractograms and refined using least square refinement method by a computer programme package Powder-X [19]. They reveal the formation of the perovskite structure with orthorhombic phase in the compound. It had been observed that the diffraction peaks shift towards the lower angle with the increase in sintering temperature. This is possibly due to the fact that the ionic radius of the Zr^{+4} ion (0.87 A°) is higher than that of Ti^{+4} ion (0.68A°) and with the substitution of Zr^{+4} ions reaction kinetics change leading to the formation of dense

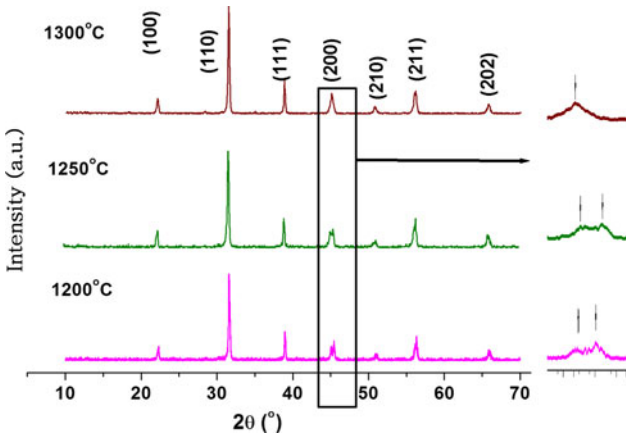


Fig. 1 X-Ray Diffractograms of $BaZr_{0.05}Ti_{0.95}O_3$ ceramics sintered at 1,200, 1,250 and 1,300 °C

ceramics at different sintering temperature [20]. It is well known that the diffusion coefficient is directly proportional to the sintering temperature and inversely proportional to the radius of the diffusing particles [21]. Hence, higher sintering temperature will be required for the diffusion of large sized Zr^{+4} ions and with the increase in sintering temperature more Zr^{+4} ions will diffuse for the same concentration of zirconium atoms and the peaks will shift towards the lower angle due to increased d spacing, as indeed observed. Moreover, the splitting of tetragonal peak (200) is observed at $2\theta = 45.262°$ in the samples sintered at 1,200 and 1,250 °C whereas in the sample sintered at 1,300 °C no splitting of any peak is observed showing the transformation towards more stable structure as the Zr substitution decreases the c-parameter and increases the a-parameter as evident by a reduced tetragonal strain (c/a) as shown in Table 1. The slow scan of the (200) peak is shown in Fig. 1, clearly indicating the splitting of this peak in the specimen sintered at lower temperature. This indicates the structure achieves more stability on sintering at higher temperature i.e. at 1,300 °C [22].

3.2 Surface morphology

Figure 2 shows the SEM micrographs of gold coated BZT ceramics sintered at temperatures 1,200, 1,250 and 1,300 °C. The samples sintered at 1,200 and 1,250 °C have

Table 1 showing the lattice parameters, tetragonal strain, average grain size, dielectric constant, Curie temperature and activation energy in ferroelectric and paraelectric region as a function of sintering temperature

Sintering Temperature (°C)	a (Å)	b (Å)	c (Å)	Tetragonal strain (c/a)	Grain size (μm)	ε'_{max}	T_c (°C)	Activation energy E_a (eV)	
								Ferroelectric	Paraelectric
1,200	3.99292	4.00363	3.99841	1.00137	0.39	411.99	125	1	0.36
1,250	3.97196	4.00122	3.99414	1.00376	0.54	646.91	115	1.04	0.72
1,300	4.02373	4.01267	4.00806	0.9961	6.15	10680.31	110	1.35	0.75

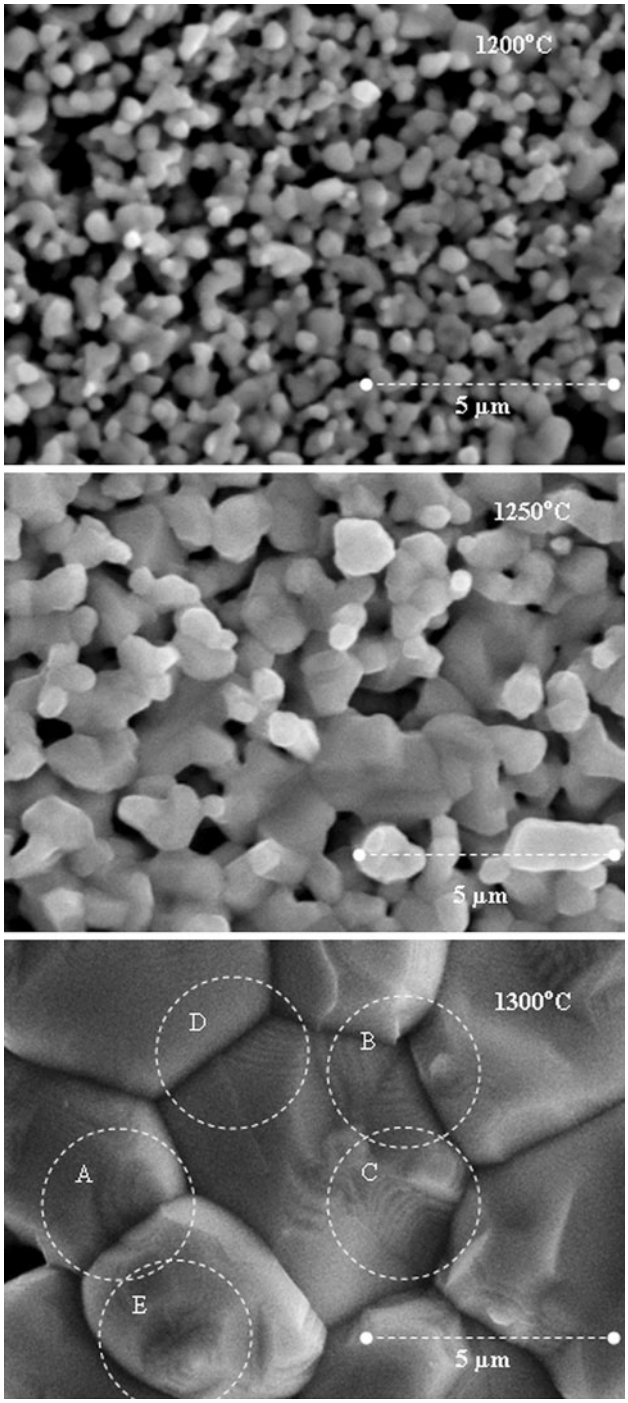


Fig. 2 SEM micrographs at various sintering temperatures

smaller grains and a porous microstructure. On increasing the sintering temperature to 1,300 °C, well developed grains with uniform grain distribution is observed. The average grain size (ASTM grain size No.) is observed to be 0.39 µm (19.3), 0.54 µm (18.4) and 6.15 µm (11.4) respectively in the samples sintered at 1,200, 1,250 and 1,300 °C. The ASTM grain size number has been calculated using equation $n = -3.3 + 6.65 \log_{10}(1/x)$ where x is average grain size

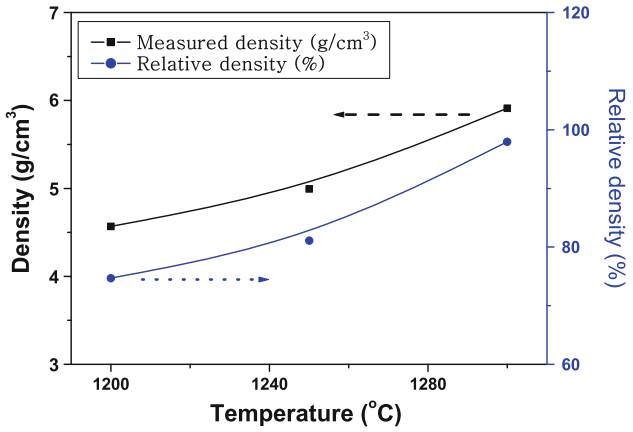


Fig. 3 Variation of density and relative density with sintering temperature

(in mm) using ASTM E112-96. The grain size in the samples increases as the sintering temperature is increased. The increase in grain size with the increase in sintering temperature can be understood as follows. The driving force for grain growth is the reduction of the total grain boundary area, which reduces the interfacial free energy for an aggregate [23]. As the interfacial energy is inversely proportional to the grain diameter [24], this causes a reduction in the total grain boundary area. The increase in the grain size reduces the porosity of the structure [25]. This process is termed as recrystallization and this is a function of temperature and time [26]. With the increase in sintering temperature from 1,200 to 1,250 and 1,300 °C, three processes recovery, recrystallization and grain growth are observed clearly in the SEM micrographs. Further, the water-mark and lamellar structures are clearly seen in the SEM micrographs as shown in Fig. 2 by the circles A, B, C, D and E. These structures appear in the high resolution SEM in these compounds which has also been earlier reported by Wei Cai et al. [20]. The characteristic feature of the domain configuration is a banded structure and the both continuous and discontinuous domains across the grain boundary have been observed in the present work as well as in the other reports [20, 27]. These domain structures influence the piezoelectric properties that are reported somewhere else.

3.3 Percent theoretical density

Figure 3 shows the measured density of the studied samples by measuring dimensions. The theoretical density of BZT ceramics was calculated to be 6.02 g/cm^3 as per JCPDS file card # 31–0174. It is observed that the experimentally measured density and relative density increases with increase in sintering temperature (Fig. 3). As seen in the SEM micrographs, the sample sintered at higher sintering temperature possess larger grains and less number of

pores. This results in the increase in density of the specimen [28].

3.4 Dielectric study

Figure 4 shows the temperature dependence of dielectric constant of the studied specimens at 1, 10 and 100 kHz. The figures show that the value of permittivity increases gradually to a maximum value (ε_m) with the increase in temperature up to the transition temperature (T_c) and then decreases gradually indicating a phase transition. A diffuse phase transition on substitution of zirconium [7] is observed. The dielectric constant increases from 412 in the sample sintered at 1,200 °C to 647 and 10,680 in the samples sintered at 1,250 and 1,300 °C respectively at 1 kHz.

3.4.1 Dielectric constant

It is observed in Fig. 4 that the dielectric constant increases with increase in sintering temperature. This can be understood as follows. At the higher sintering temperature the average grain size increases making the domain wall motion easier resulting in an increased dielectric constant [29]. As the sintering is carried out in air, the lattice undergoes reoxidation and it forms a strong insulating layer around a grain with a higher bulk resistance. It results in high interfacial polarization leading to the increase in dielectric constant [30]; as the grain boundaries have large resistance and capacitance as compared to the bulk [31]. Moreover, the difference in the conductivity of the bulk and grain boundary results in an increase in surface charge accumulation and this in turn, results in the increase of interfacial polarization. With the increase in sintering temperature from 1,200 to 1,300 °C, relative density of the samples increased from 74 to 98 % indicating the grain growth and is responsible for a large increase in dielectric constant from 412 to 10680.

It is also observed in Fig. 4 that T_c reduces from 125 to 120 and 115 °C on substitution of zirconium as the sintering temperature is increased. This can then be possibly due to the weak bonding force between the Ti^{+4} and the oxygen ion of the ABO_3 perovskite structure. As the Zr–O bonds (776.1 kJ/mol) are stronger than Ti–O bonds (672.4 kJ/mol), hence the Zr^{+4} ions can be substituted at lower temperature for cubic phase formation, hence the phase transition temperature is reduced [32]. Also, the weakening of Ti–O bond leads to a weaker distortion of the octahedron and the substitution of Zr might induce a "break" of the Ti–O chains; this could bring about a decrease in the c/a ratio as depicted in Table 1. This "break" is responsible for the drop in Curie temperature

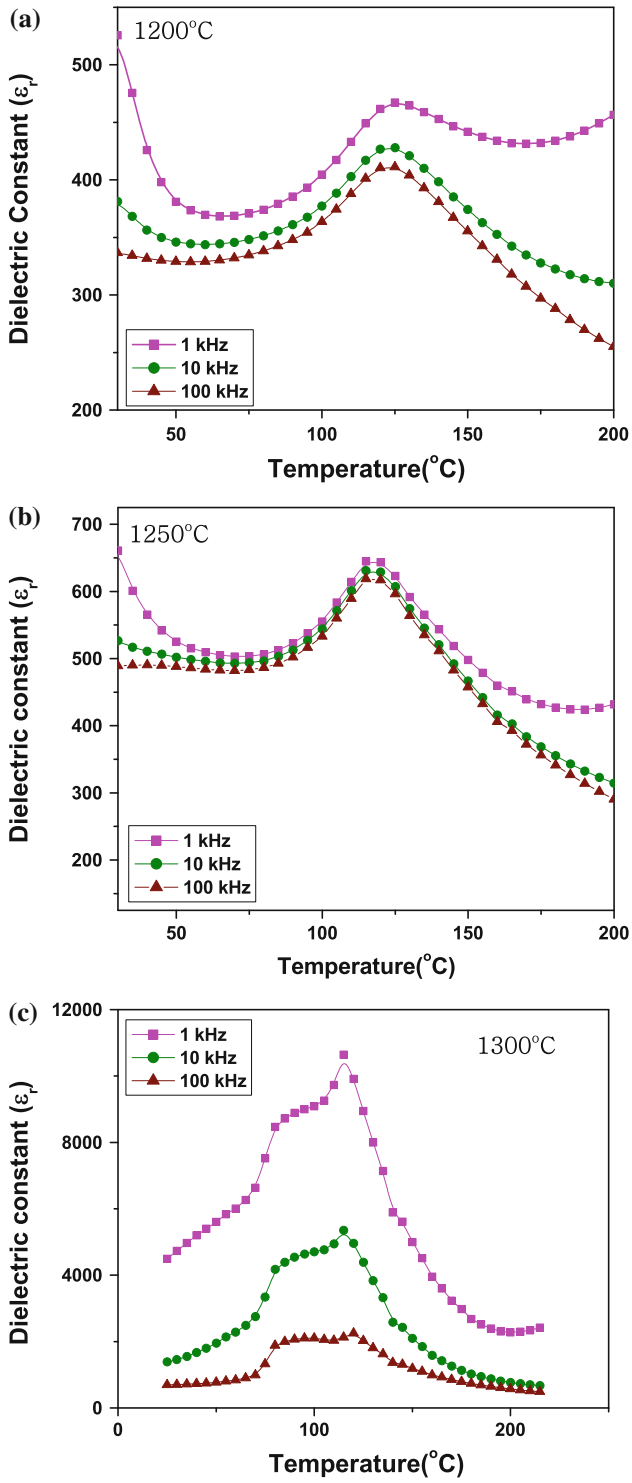


Fig. 4 Variation of dielectric constant with temperature at frequencies 1 kHz, 10 kHz and 100 kHz **a** 1,200 °C **b** 1,250 °C **c** 1,300 °C

[33, 34]. It is also observed in Fig. 4 that there is one more phase transition, other than at T_c, which is attributed to the tetragonal to orthorhombic phase transition [34].

3.4.2 Dielectric loss

The dielectric loss of the studied samples from room temperature to 300 °C is shown in Fig. 5. The loss decreases with the increase in temperature and is nearly independent of temperature up to about 200 °C and then increases rapidly with increase in temperature. This sharp increase of dielectric loss in the high temperature region may be attributed to the increased mobility of charge carriers arising from defects or vacancies in the sample [35]. The defect chemical reaction of Zr^{+4} substitution for Ti^{+4} is represented as [22]:

$$Zr^{+4} \rightarrow Zr_{Ti}^{*} + 2O_{O}^{*}$$

$$Ti^{+4} + e' \rightarrow Ti^{+3}$$

At the higher sintering temperature, Ti^{+4} reduce to Ti^{+3} and there is electron hopping between Ti^{+4} and Ti^{+3} ions. It is observed that the loss peak is formed at Curie temperature showing the phase transition. At some suitable frequencies of the applied alternating field, resonance in dielectric behavior is often observed and this peak may be due to the resonance achieved [36].

3.4.3 DC conductivity measurements

The effect of the temperature on the dc electrical conductivity σ of the studied specimens is shown in Fig. 6. In all the samples, it is observed that conductivity decreases with the increase in temperature nearly up to Curie temperature and then it increases subsequently. The decrease in conductivity is due to decrease in the drift mobility of the charge carriers [37]. Table 1 gives the values of activation energies (E_a) calculated for the studied samples, using Fig. 6 and the Arrhenius equation:

$$\sigma_{dc} = \sigma_o \exp\left(\frac{-E_a}{k_B T}\right)$$

where k_B is the Boltzmann constant. It is found that E_a is higher in the specimen sintered at 1,300 °C than that in the sample sintered at 1,200 °C. It is known that the specimen with a smaller grain size contains higher number of grain boundaries which results in higher contribution of grain boundary resistance towards the overall resistance [38]. The values of activation energy in the ferroelectric and paraelectric region in the studied specimens are listed in Table 1. The observed existence of two activation energies may be associated with the structural changes taking place at transition temperature [39]. It is well known that conductivity in ferroelectrics is a complex phenomenon. It is affected by the presence of defects, vacancies, imperfections and changes in lattice parameters [38]. These materials are sintered at higher sintering

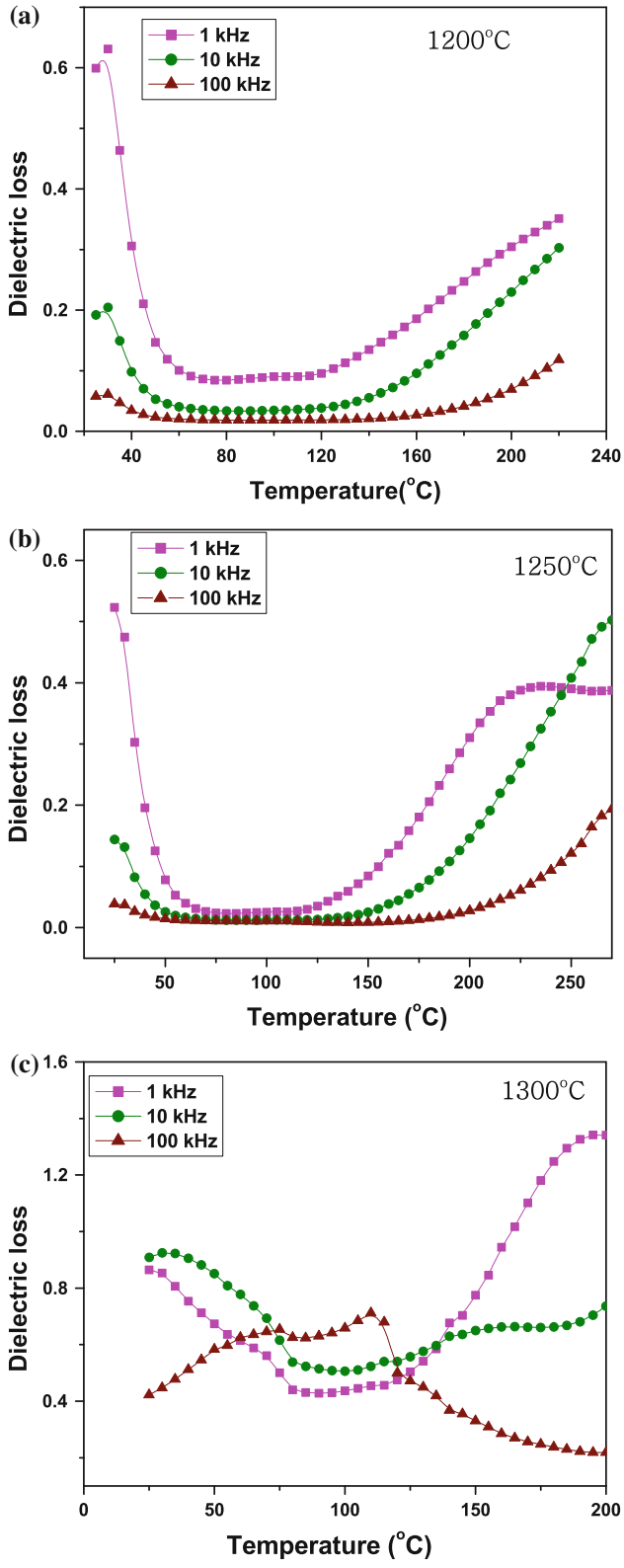


Fig. 5 Variation of dielectric loss (tan δ) with temperature at frequencies 1, 10 and 100 kHz **a** 1,200 °C **b** 1,250 °C **c** 1,300 °C

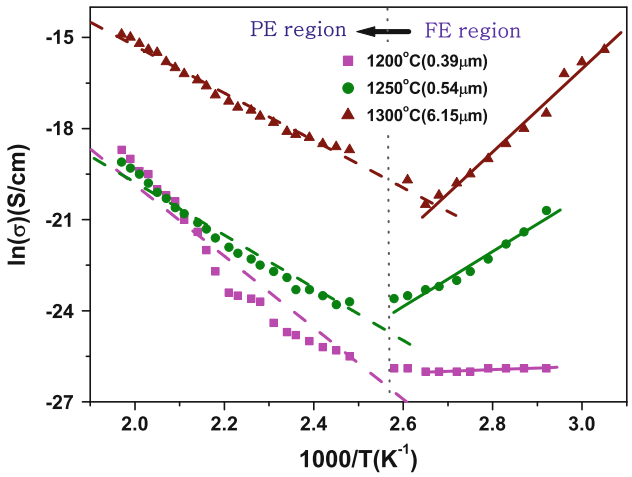


Fig. 6 Variation of dc electrical conductivity σ with temperature for the sintering temperature

temperature where they are expected to lose traces of oxygen as per the following representation:

$$O_O \leftrightarrow \frac{1}{2}O_2 + V_O^{\bullet\bullet} + 2e'$$

where all the species are written in accordance with Kroger–Vink equation of defects. The conduction can be due to the singly and doubly charged oxygen vacancies [40]. For the formation of singly charged oxygen vacancies, the value of activation energy lies in the range of 0.2–0.4 eV and for the doubly charged oxygen vacancies the value of activation energy lies in between 0.8 and 1.0 eV [40]. As seen in the Table 1, the values of activation energy lie in between 0.7 and 1.0 eV, hence the observed conductivity variations may be attributed to the doubly charged oxygen vacancies in the specimens.

3.4.4 Diffuse phase transition

A diffuse phase transition is generally characterized by: (a) broadening in the dielectric constant (ε) versus temperature (T) curve; (b) a relatively large separation (in temperature) between the maximum of the real (dielectric constant) and imaginary (dielectric loss) parts of the dielectric spectrum; (c) a deviation from Curie–Weiss law in the vicinity of T_m and (d) frequency dispersion of both ε and tan δ in the transition region thereby implying a frequency dependence of T_m [12].

There is broad dielectric constant temperature curve in the transition temperature region, as seen in Fig. 4, which can be attributed to inhomogeneous distribution of Zr ions at the Ti sites [41]. It can also be attributed to the tolerance factor of the perovskite structure. The strain in the structure is characterized by Goldschmidt tolerance factor [42]. For perovskites with the general formula of ABO_3, the following equation is used to calculate the tolerance factor, t:

$$t = \frac{R_O + R_A}{\sqrt{2}(R_B + R_O)}$$

where R_A is the radius of A-site atom, R_B is the radius of B-site atom and R_O is the radius of the O atom.

For the composition in the present work, the above equation can be rewritten as:

$$t = \frac{R_O + R_{Ba}}{\sqrt{2}(xR_{Zr} + (1-x)\,R_{Ti} + R_O)}$$

For a stable perovskite structure, the value of tolerance factor (t) should be close to unity. In case of $BaTiO_3$ value of t is greater than unity and hence the normal ferroelectric phase is stable [43] and the transition is a sharp phase transition. Zr^{+4} ion substitution at B-sites reduces the value of tolerance factor to 0.96, as calculated from the above mentioned relation and it induces paraelectric behavior. The reduction in the value of tolerance factor is attributed to the higher ionic diameter of Zr^{+4} ion in comparison to Ti^{+4} ion. This introduces disorderedness in the system leading to the diffuse phase transition.

A modified Curie–Weiss law [44] has been proposed to describe the diffuseness of a phase transition

$$1/\varepsilon - 1/\varepsilon_m = (T - T_m)^{\gamma}/C',$$

where, ε is relative dielectric constant at Temperature T and ε_m is maximum relative dielectric constant at temperature T_m, C' is Curie constant. The parameter γ is regarded as a measure of the diffuseness of the phase transition: $\gamma = 1$ corresponds to the normal ferroelectric behavior while $\gamma = 2$ corresponds to the diffuse phase transition. The plots of ln $(1/\varepsilon - 1/\varepsilon_m)$ as a function of ln $(T - T_m)$ for the different sintering temperatures (1,200, 1,250, 1,300 °C) at different frequencies (1, 10, 100 kHz) are shown in Fig. 7. The crystal shows disordered phase transition due to large size of Zr^{+4} ions [45].

3.5 Hysteresis loops

Figure 8 shows the hysteresis loops of the studied specimens measured at room temperature. Unsaturated loops are observed in the sample sintered at low temperature leading to poor ferroelectric characteristics [46]. However, a proper saturated loop is observed in the sample sintered at 1,300 °C. The remanent polarization ($2P_r$) increases to 22 μC/cm^2 in the sample sintered at 1,300 °C. There are two major effects of grain boundary on polarization. In the samples with lower grain size, there is large number of grain boundaries [20] and it is well known that grain boundary is a low permittivity region; hence these samples should show poor ferroelectric

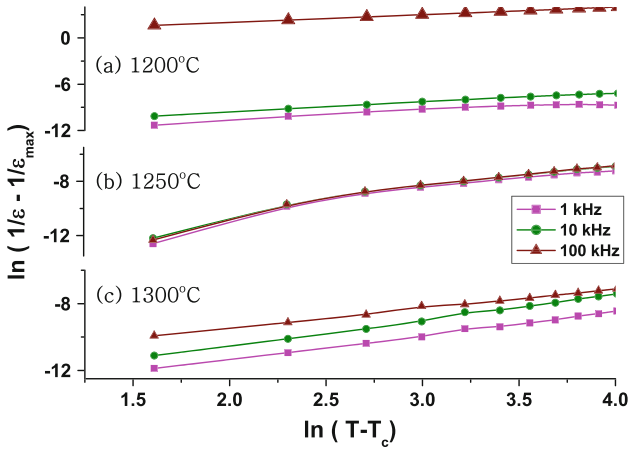


Fig. 7 The plot of ln $(1/\varepsilon–1/\varepsilon_m)$ as a function of ln $(T - T_m)$ for the sintering temperature at different frequencies

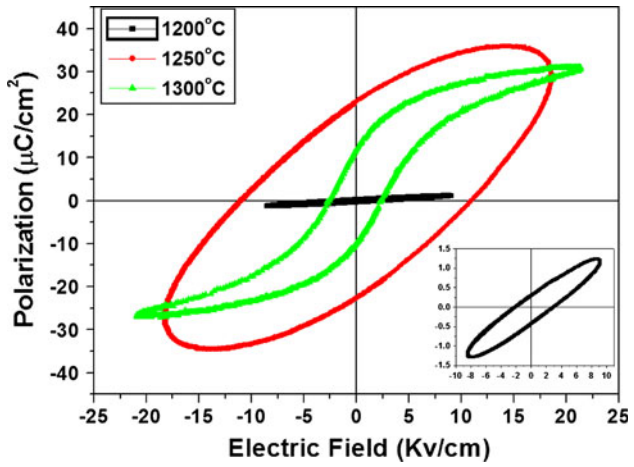


Fig. 8 Ferroelectric hysteresis loops showing remanent polarization (P_r) and coercive field (E_c) at (i) 1,200 °C (ii) 1,250 °C (iii) 1,300 °C

characteristics. Also, it is well known that there is polarization discontinuity between grain boundary and grain surface which reduces the remanent polarization. Both these factors deteriorate the ferroelectric characteristics, as indeed observed in the present work. It is also observed in Fig. 7 that the coercive field decreases with increase in sintering temperature. This might have reduced the breakdown strength. This may be due to the fact that energy barrier for switching ferroelectric domain gets reduced as the grain size increases. Polarization reversal of a ferroelectric domain is much easier inside a larger grain compared to that in a smaller grain [43, 47].

4 Conclusions

The synthesized compounds are formed with perovskite structure in orthorhombic phase. As the sintering temperature increased from 1,200 to 1,300 °C, the average grain size increased from 0.39 to 6.15 μm. The measured values of density of the samples increase with increase in sintering temperature. The SEM micrographs show domain structures in the sample sintered at 1,300 °C. The sample sintered at 1,300 °C show the highest value of dielectric constant and remanent polarization. In the present work, the results show that the sample sintered at 1,300 °C possesses improved dielectric and ferroelectric characteristics.

References

1. S.M. Neirman, The Curie point temperature of $Ba(Ti_{1-x}Zr_x)O_3$ solid solutions. J. Mater. Sci. **23**, 3973–3980 (1988)
2. Y. Sakabe, K. Minai, K. Wakino, High-dielectric constant ceramics for base metal monolithic capacitor. Jpn. J. Appl. Phys. **20**(Supplement 20–4), 147–150 (1981)
3. H. Nemoto, I. Oda, Direct examinations of PTC action of single grain boundaries in semiconducting $BaTiO_3$ Ceramics. J. Am. Ceram. Soc. **63**, 398 (1980)
4. M. Kuwabara, K. Morimo, T. Matsunaga, Single grain boundaries in PTC resistors. J. Am. Ceram. Soc. **79**, 997 (1996)
5. M. McQuarrie, F.W. Behnke, Structural and dielectric studies in the system (Ba, Ca) (Ti, Zr) O_3. J. Am. Ceram. Soc. **37**, 539 (1954)
6. A. Dixit, S.B. Majumder, R.S. Katiyar, A.S. Bhalla, Studies on the relaxor behaviour of sol-gel derived Ba $(Zr_x\ Ti_{1-x})$ O3 $(0.30 \leq x \leq 0.70)$ thin films. J. Mater. Sci. **41**, 87–96 (2006)
7. D. Hennings, A. Schnell, G. Simon, Diffuse ferroelectric phase transitions in $BaTi_{1-y}\ Zr_y\ O_3$. J. Am. Ceram. Soc. **65**, 539–544 (1982)
8. Y. Zhi, A. Chen, R. Guo, A.S. Bhalla, Piezoelectric and strain properties of $Ba(Ti_{1-x}Zr_x)O_3$ ceramics. J. Appl. Phys. **92**, 1489–1493 (2002)
9. R. C. Kell, N. J. Hellicar, Structural transformations in transformer materials of barium-titanate-zirconate. Acustica **6**, 235–238 (1956)
10. R. Farhi, M.E.I. Marssi, A. Simon, J. Ravez, A Raman and dielectric study of ferroelectric $Ba(Ti_{1-x}Zr_x)O_3$ ceramics. Eur. Phys. J. B **9**, 599–604 (1999)
11. Y. Zhi, A. Chen, R. Guo, A.S. Bhalla, Ferroelectric-relaxor behavior of $Ba(Ti_{0.7}Zr_{0.3})O_3$ ceramics. J. Appl. Phys. **92**, 2655–2657 (2002)
12. S. K. Rout, Ph. D. Thesis, Phase formation and dielectric studies of some $BaO\text{-}TiO_2\text{-}ZrO_2$ based perovskite system. (2006)
13. M. Deri, *Ferroelectric Ceramics* (Gordon and Breach, New York, 1969)
14. B. Jaffe, W. Cook, H. Jaffe, *Piezoelectric Ceramics* (Academic Press, London, 1971)
15. E. Brajer, Polycrystalline ceramic material. US Patent 2708243: 1955
16. X.G. Tang, K.H. Chew, H.L.W. Chan, Diffuse phase transition and dielectric tunability of $BaTi_{1-y}\ Zr_y\ O_3$ relaxor ferroelectric ceramics. Acta Mater. **52**, 5177–5183 (2004)
17. C. Heremans, H.L. Tuller, Field-induced antiferroelectric-ferroelectric phase transitions in the $Pb_{0.98}\ La_{0.02}\ (Zr_{0.70}\ Hf_{0.30})_{1-x}Ti_xO_3$ system. J. Appl. Phys. **87**, 1458–1486 (2000)
18. E. F. Alberta, P. W. Rehrig, W. S. Hackenberger, C. A. Randall, T. R. Shrout, High temperature morphotrophic phase boundary piezoelectrics, IEEE Ultrasonics Symposium 1991–1994 (2003). doi:10.1109/ULTSYM.2003.1293307
19. C. Dong, Powder-X: Windows-95-based program for powder X-ray diffraction data processing. 32:838

20. W. Cai, C. Fu, J. Gao, H. Chen, Effect of grain size on domain structure and ferroelectric properties of barium zirconate titanate ceramics. J Alloys Compd **480**, 870–873 (2009)
21. K. Huang, *Statistical Mechanics* (Wiley, New York, 1987)
22. S.W. Zhang, H. Zhang, B.P. Zhang, G. Zhao, Dielectric and piezoelectric properties of characteristics of $(Ba_{0.95}Ca_{0.05})(Ti_{0.88}Zr_{0.12})O_3$ ceramics sintered in a protective atmosphere. J Eur Ceram Soc **29**, 3235–3242 (2009)
23. W.D. Kingrey, H.K. Bowen, D.R. Uhlmann, *Introduction to Ceramics*, 2nd edn. (Wiley, USA, 1976)
24. J. Tullis, R.A. Yund, Grain growth kinetics of quartz and calcite aggregates. J Geol **90-3**, 301–308 (1982)
25. M.W. Barsoum, *Fundamentals of Ceramics* (McGraw Hill Companies, New York, 1996)
26. W.D. Callister, D.G. Rethwisch, *Materials Science and Engineering: An Introduction* (Wiley, USA, 2010)
27. S. Tsurekawa, K. Ibaraki, K. Kawahara, T. Watanabe, The continuity of ferroelectric domains at grain boundaries in lead zirconate titanate. Scripta Materialia **56**, 577–580 (2007)
28. K. Ramam, M. Lopez, Barium modified lead lanthanum strontium zirconium niobium titanate for dielectric and piezoelectric properties. J. Eur. Ceram. Soc. **27**, 3141–3147 (2007)
29. S.B. Reddy, K.P. Rao, M.S.R. Rao, Observation of high permittivity in Ho substituted $BaZr_{0.1}Ti_{0.9}O_3$ ceramics. Appl. Phys. Lett. **91**, 022917 (2007)
30. D.C. Sinclair, A.R. West, Impedance and modulus spectroscopy of semiconducting $BaTiO_3$ showing positive temperature coefficient of resistance. J. Appl. Phys. **66**, 3850 (1989)
31. D.C. Sinclair, T.B. Adams, F.D. Morrison, A.R. West, $CaCu_3Ti_4O_{12}$: one-step internal barrier layer capacitor Appl. Phys. Lett. **80**, 2153 (2002)
32. S.G. Lee, D.S. Kan, Dielectric properties of ZrO_2 doped (Ba, Sr, Ca) TiO_3 ceramics for tunable microwave device applications. Mater. Lett. **57**, 1629 (2002)
33. W.R. Huo, Y.F. Qu, Effects of $Bi_{1/2}Na_{1/2}TiO_3$ on the Curie temperature and the PTC effects of $BaTiO_3$-based positive temperature coefficient ceramics. Sens. Actuators, A **128**, 265 (2006)
34. N. Sawangwan, J. Barrel, K. Mackenzie, T. Tunaksiri, The effect of Zr content on electrical properties of $Ba(Ti_{1-x}Zr_x)O_3$ ceramics. Appl. Phys. A **90**, 723 (2008)
35. P. Ganguly, A.K. Jha, Structural and electrical properties of $Ba_{5-x}Ca_xSmTi_3Nb_7O_{30}$ (x = 0–5) ferroelectric ceramics. J. Alloys Compd **495**, 7–12 (2010)
36. R.C. Buchanan, *Ceramic Materials for Electronics* (Marcel Dekker, New York, 1991)
37. N. Ponpandian, A. Narayanasamy, Influence of grain size and structural changes on the electrical properties of nanocrystalline zinc ferrite. J. Appl. Phys. **92**, 2770 (2002)
38. S.A. Fadnavis, A.G. Katpatal, Anomalous dc conductance in Nd^{+3}-doped ferroelectric lead germanate single crystal. Bull. Mater. Sci. **20**, 1097 (1997)
39. P. Ganguly, A.K. Jha, Enhanced characteristics of $Ba_5SmTi_3Nb_7O_{30}$ ferroelectric nanocrystalline ceramic prepared by mechanical activation process: a comparative study. Mater. Res. Bull. **46**, 692 (2011)
40. R.K. Dwivedi, D. Kumar, O. Prakash, Valence compensated perovskite oxide system $Ca_{1-x}La_xTi_{1-x}Cr_xO_3$. J. Mater. Sci. **36**, 3641–3648 (2001)
41. U. Weber, G. Greuel, U. Boettger, S. Weber, D. Hennings, R. Waser, Dielectric properties of Ba(Ti, Zr)O_3-based ferroelectrics for capacitor applications. J. Am. Ceram. Soc. **84**, 759 (2001)
42. V. M. Goldschmidt, Shrifter Nofke Videnskaps-Akad. Oslo I: mat-Naturv. Kl. No. 2, 8 (1926)
43. H.T. Martirena, J.C. Burfoot, Grain size effects on properties of some ferroelectric ceramics. J. Phys. C: Solid State Phys. **7**, 3182 (1974)
44. G. Burns, Dirty displacive ferroelectrics. Phys. Rev. B **13**, 215 (1976)
45. M. Kumar, A. Garg, R. Kumar, M.C. Bhatnagar, Structural, dielectric and ferroelectric studies of $Ba_{0.9}Sr_{0.1}Zr_xTi_{1-x}O_3$ ceramics prepared by the sol-gel method. Phys. B **403**, 1819 (2008)
46. N. Nanakorn, P. Jalupoom, N. Vaneesorn, A. Thanaboonsombut, Dielectric and ferroelectric properties of $BaTi_{1-x}$ Zr_x O_3 ferroelectric properties. Ceram. Int. **34**, 779–782 (2008)
47. P.A. Jha, A.K. Jha, Influence of processing conditions on the grain growth and electrical properties of barium zirconate titanate ferroelectric ceramics. J. Alloys Compd **513**, 580–585 (2012)

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 8, October 2012

Evaluating NIST Metric for English to Hindi Language Using ManTra Machine Translation Engine

Neeraj Tomer[1]
AIM & ACT
Banasthali University Banasthali
Jaipur, India

Deepa Sinha[2]
Department of Mathematics
South Asian University
New Delhi, India

Abstract: Evaluation of MT is required for Indian languages because the same MT is not works in Indian language as in European languages due to the language structure. So, there is a great need to develop appropriate evaluation metric for the Indian language MT. The present research work aims at studying the Evaluation of Machine Translation Evaluation's NIST metric for English to Hindi for tourism domain using the output of ManTra, a translation system. Machine Translation Evaluation has been widely recognized by the Machine Translation community. The main objective of MT is to break the language barrier in a multilingual nation like India.

Keywords: MTE- Machine Translation Evaluation, MT – Machine Translation, EILMT –Evaluation of Indian Language Machine Translation, ManTra – MAchiNe Assisted TRAnslation Technology, Tr – Tourism

INTRODUCTION

Indian languages are highly inflectional, with a rich morphology, relatively free word order, and default sentence structure as Subject-Object-Verb. In addition, there are many stylistic differences. So the evaluation of MT is required for Indian languages because the same MT is not works in Indian language as in European languages. The same tools are not used directly because of the language structure. So, there is a great need to develop appropriate evaluation metric for the Indian language MT.

English is understood by less than 3% of Indian population. Hindi, which is official language of the country, is used by more than 400 million people. MT assumes a much greater significance in breaking the language barrier within the country's sociological structure. The main objective of MT is to break the language barrier in a multilingual nation like India. English is a highly positional language with rudimentary morphology, and default sentence structure as Subject-Verb-Object. The present research work aims at studying the "Evaluation of Machine Translation Evaluation's NIST Metric for English to Hindi" for tourism domain. The present research work is the study of statistical evaluation of machine translation evaluation for English to Hindi. The research aims to study the correlation between automatic and human assessment of MT quality for English to Hindi. The main goal of our experiment is to determine how well a variety of automatic evaluation metric correlated with human judgment.

In the present work we propose to work with corpora in the tourism domain and limit the study to English – Hindi language pair. It may be assumed that the inferences drawn from the results will be largely applicable to translation for English to other Indian Languages. Our test data consisted of a set of English sentences that have been translated from expert and non-expert translators. The English source sentences were randomly selected from the corpus of tourism domain. These sentences are taken randomly from the different resources like websites, pamphlets etc. Each output sentence was score by Hindi speaking human evaluators who were also familiar with English. It may be assumed that the inferences drawn from the results will be largely applicable to translation for English to other Indian Languages, as assumption which will have to be tested for validity. We intend to be consider the following MT engine in our study-

ManTra: C-DAC Pune has developed a translation system called ManTra. The work in ManTra has to be viewed in its potentiality of translating the bulk of texts produced in daily official activities. The system is facilitated with pre-processing and post-processing tools, which enables the user to overcome the problems/errors with minimum effort. The strategy used for translation is: NOT Word to Word; NOR Rule to Rule; BUT Lexical Tree to Lexical Tree.

OBJECTIVE

The main goal of this work is to determine how well a variety of automatic evaluation metrics correlated with human scores. The other specific objectives of the present work are as follows.

1. To design and develop the parallel corpora for deployment in automatic evaluation of English to Hindi machine translation systems.
2. Assessing how good the existing automatic evaluation metrics NIST, will be as MT evaluating strategy for evaluation of Indian language machine translation systems by comparing the results obtained by this with human evaluator's scores by correlation study.
3. To study the statistical significance of the evaluation results as above, in particular the effect of-
   - size of corpus
   - sample size variations
   - increase in number of reference translations

Creation of parallel corpora: Corpus quality plays a significant role in automatic evaluation. Automatic metrics can be expected to correlate very highly with human judgments only if the reference texts used are of high quality, or rather, can be expected to be judged high quality by the human evaluators. The procedure for creation of parallel corpora is as under:

1. Collect English corpus from the domain from various resources.
2. Generate multiple references (we limit it to three) for each sentence by getting the source sentence translated by different expert translators.

3. XMLise the source and translated references for use in Automatic evaluation

Description of Corpus

Domain	Source Language	Target Language	No. of Sentences	No. of Human Translation	Name of MT Engine
Tourism	English	Hindi	1000	3	Mantra

For the corpus collection our first motive was to collect as possible to get better translation quality and a wide range vocabulary. For this purpose the first corpus we selected to use in our study is collected from different sources. We have manually aligned the sentence pairs.

In our study for tourism domain we take 1000 sentences. When the text has been collected, we distributed this collected text in the form of Word File. Each word files having the 100 sentences of the particular domain. In this work our calculation will be based on four files- source file and three reference files. Reference files are translated by the language experts. We give the file a different identification. For e.g. our first file name is Tr_0001_En where Tr_ for tourism 0001 means this is the first file and En means this is the Candidate file. We treat this as the candidate file. In the same way our identification for the Hindi File is Tr_0001_Hi, in this Hi is for the Hindi file and we have called this a reference file. As we already mention that we are taking the three references we named them reference 1(R1), reference 2(R2), reference 3(R3). In the study we take the candidate sentence and the reference sentences, as shown below. For e.g.

Source Sentence: Guru Shikhar is the highest peak on the Mount Abu which provides an excellent view of the whole town.

Candidate Sentence: पहाड़ी अबु जो प्रदान सर्वोत्तम दृश्य का सभी नगर पर गुरु शिखहर उच्चतम शिखर

Reference Sentences:

R1: गुरूशिखर माऊण्ट आबू की सबसे ऊँची चोटी है जहां से पूरे शहर का उत्तम दृश्य देखने को मिलता है।

R2: गुरू शिखर माउण्ट आबू की सबसे ऊंची चोटी है जो पूरे शहर का अत्यंत उत्कृष्ट दृश्य प्रदान करती है।

R3: गुरू शिखर माउण्ट आबू पर सबसे ऊँचा शिखर है जो पूरे शहर का उत्कृष्ट दृश्य प्रदान करता है।

HUMAN EVALUATION

Human evaluation is always best choice for the evaluation of MT but it is impractical in many cases, since it might take weeks or even months (though the results are required within days). It is also costly, due to the necessity of having a well trained personnel who is fluent in both the languages, source and targeted. While using human evaluation one should take care for maintaining objectivity. Due to these problems, interest in automatic evaluation has grown in recent years. Every sentence was assigned a grade in accordance with the four point scale for adequacy.

AUTOMATIC EVALUATION BY NIST METRIC

We used NIST evaluation metric for this study. This metric is specially designed for English to Hindi. NIST metric, designed for evaluating MT quality, scores candidate sentences by counting the number of n-gram matches between candidate and reference sentences. NIST metric is probably known as the best known automatic evaluation for MT. To check how close a candidate translation is to a reference translation, an n-gram comparison is done between both. Metric is designed from matching of candidate translation and reference translations. We have chosen correlation analysis to evaluate the similarity between automatic MT evaluations and human evaluation. Next, we obtain scores of evaluation of every translated sentence from both MT engines. The outputs from both MT systems were scored by human judges. We used this human scoring as the benchmark to judge the automatic evaluations. The same MT output was then evaluated using both the automatic scoring systems. The automatically scored segments were analyzed for Spearman's Rank Correlation with the ranking defined by the categorical scores assigned by the human judges. Increases in correlation indicate that the automatic systems are more similar to a human in ranking the MT output. Statistical significance is an estimate of the degree, to which the true translation quality lays within a confidence interval around the measurement on the test sets. A commonly used level of reliability of the result is 95%. To reach at decision, we have to set up a hypothesis and compute p-value to get final conclusion.

The present research is the study of statistical evaluation of machine translation evaluation's NIST metric. The research aims to study the correlation between automatic and human assessment of MT quality for English to Hindi. While most studies report the correlation between human evaluation and automatic evaluation at corpus level, our study examines their correlation at sentence level. The focus in this work is to examine the correlation between human evaluation and automatic evaluation and its significance value, not to discuss the translation quality. In short we can say that this research is the study of statistical significance of the evaluated results, in particular the effect of sample size variations.

So, firstly we take source sentences and then get these sentences translated by our MT engine, here we consider the Anuvadaksh. We have the different references of these sentences. After doing this we do the evaluations of these sentences human as well as the automatic evaluations and we collect the individual scores of the given sentences considering all the three references one by one. The following table shows the individual scores of the five sentences (particular sentences can be seen at the end of the paper) using different no. of references.

Table 1: Human Evaluation and NIST Evaluation scores

S. No.	NIST Score			
	Human Eval.	one no. of reference	two no. of references	three no. of references
1.	0.75	0	0.0792	0.0792
2.	0.5	0.2221	0.2511	0.2511
3.	0.75	0.0695	0.1508	0.1508
4.	0.75	0.1464	0.1797	0.1797
5.	0.75	0	0.1394	0.1394

In this way we also collect the individual scores of all the sample sizes like 20, 60,100,200,300,500 and 1000 sentences. After this we do the correlation analysis of these values. In order to calculate the correlation with human judgements during evaluation, we use all English–Hindi human rankings distributed during this shared evaluation task for estimating the correlation of automatic metrics to human judgements of translation quality, were used for our experiments. In our study the rank is provided at the sentence level.

For correlation analysis we calculate the correlation between human evaluation and automatic evaluations one by one by the Spearman's Rank Correlation method. The Spearman's rank correlation coefficient is given as (when ranks are not repeated)-

$$\rho = 1 - \left(\frac{6\sum_{i=1}^{n} d^2}{n(n^2-1)} \right)$$

where d is the difference between corresponding values in rankings and n is the length of the rankings. An automatic evaluation metric with a higher correlation value is considered to make predictions that are more similar to the human judgements than a metric with a lower value. Firstly, we calculate the correlation value in between the human evaluation and automatic evaluation NIST metric means human evaluation with NIST for sample size 20, 60, 100, 200, 300, 500 and 1000.

Table 2: Correlation (ρ) values

Sample Size	ρ values		
	one no. of reference	two no. of references	three no. of references
20	.110	.360	.360
60	-.075	.017	.017
100	.060	.048	.048
200	.407	.287	.407
300	.376	.274	.274
500	.304	.245	.245
1000	.274	.245	.274

After calculating the correlation, we need to find out which type of correlation is there between the variables and of which degree and whether the values of the correlation are significant.

ANALYSIS OF STATISTICAL SIGNIFICANCE TEST FOR HUMAN EVALUATION AND AUTOMATIC EVALUATION

Statistical significance is an estimate of the degree, to which the true translation quality lays within a confidence interval around the measurement on the test sets. A commonly used level of reliability of the result is 95%, for e.g. if, say, 100 sentence translations are evaluated, and 30 are found correct, what can we say about the true translation quality of the system? To reach at decision, we have to set up a hypothesis and compute p-value to get final conclusion that whether there is any correlation between the human evaluations and automatic evaluations. If yes, then what is the type and degree of correlation? Also what is the significance of the correlation value? In this work we set the hypothesis that there is no correlation between the values of human and automatic evaluation. The p-value will provide the answer about the significance of the correlation value.

A Z-test is a statistical test for which the distribution of the test statistic under the null hypothesis can be approximated by a normal distribution. For each significance level, the Z-test has a single critical value (for example, 1.96 for 5% two tailed) which makes it more convenient than the Student's t-test which has separate critical values for each sample size. The test statistic is calculated as:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the sample means, s_1^2 and s_2^2 are the sample variances, n_1 and n_2 are the sample sizes and z is a quartile from the standard normal distribution.

Table 3: p-values of output of Anuvadaksh using different no. of references

Sample Size	p-values		
	one no. of reference	two no. of references	three no. of references
20	0.0001	0.0001	0.0001
60	0.0001	0.0001	0.0001
100	0.0001	0.0001	0.0001
200	0.1977	0.1814	0.1977
300	0.2358	0.2033	0.2033
500	0.2643	0.2266	0.2266
1000	0.2451	0.2266	0.2451

Now on the basis of these values we conclude our results like which type and degree of correlation is there between the given variables and whether the correlation results are significant. In the above example we have done all the calculations by considering the single reference sentence and in tourism domain using 5 numbers of sentences.

But in our research work we consider the different references like 1, 2, 3 and we use the different sample sizes like 20, 60, 100, 200, 300, 500, and 1000. We see whether the results remains uniform for different sample sizes and different number of references in particular domains.

For above calculation we used following sentences:
English Sentences:

1. The best way to experience the real magic of Thar Desert, Rajasthan is with the help of Desert Safari.
2. Camel safari could be the best choice for the adventure-seeking tourists moving around the interiors of Thar Desert, witnessing the cities and historical ruins of the city.
3. Sam Sand Dunes is a perfect tourist destination known for the sunset and sunrise point.
4. Manvar Desert is best for the outdoor adventures and excitement.
5. Bikaner is located to the north of Rajasthan popularly known as a camel country.

Candidate Sentences (translated by ManTra):

1. महसूस करने के लिए योग्यता यात्रा का थार मरूस्थल , रजस्थन के साथ सहायता का सर्वोत्तम तरीका वास्तविक जादू
2. ऊँट यात्रा सबसे अच्छा पसंद साहस प्राप्त पर्यटकों के लिए थार मरूस्थल का आंतरिक करीब , सुधार कर रही हैं नगर का शहरों और ऐतिहासिक खंडहर देखरहे हैं
3. सम बालू बालू के टीले पूरा पर्यटक मंजिल प्रसिद्ध के लिए सूर्यास्त और सूर्योदय इशारा करते हैं
4. मांवर योग्यता सबसे अच्छा के लिए बाहरी साहसिक कार्यों और उत्तेजना
5. रजस्थन सामान्यतया प्रसिद्ध के रूप में ऊँट देश का बीकानेर स्थित तक उत्तर

RESULTS

In the domain tourism there is significance difference between the average evaluation score of human with NIST at 5% level of significance and this is for sample sizes 20, 60 and 100.

In Table 2 (Correlation (ρ) values) correlation value for NIST is .110 and .360 these values are for sample size 20 and for one and two number of references which is significant at 5% level of significance. A similar result is seen in the case of sample size 100 for all three references. But for the sample sizes 200, 300, 500 and 1000 value of correlation is insignificant on the given level of significance.

CONCLUSION

Corpus quality plays a significant role in automatic evaluation. Automatic metrics can be expected to correlate highly with human judgments only if the reference texts used are of high quality. This work will help to give the feedback of the MT engines. In this way we may make the changes in the MT engines and further we may revise the study.

ACKNOWLEDGMENT

The present research work was carried under the research project "English to Indian Languages Machine Translation System (EILMT)", sponsored by TDIL, Ministry of Communications and Information Technology, Government of India. With stupendous ecstasy and profundity of complacency, we pronounce utmost of gratitude to Late Prof. Rekha Govil, Vice Chancellor, Jyoti Vidyapith, Jaipur Rajasthan.

REFERENCES

1. Akiba, Yasuhir Taro Watanabe, Eiichiro Sumita, (2002): "Using Language and Translation Models to Select the Best among Outputs from Multiple {MT} System", Proceeding of Colong, 8-14.
2. Alex Kulesza, Stuart M Shieber (2004): "A Learning Approach to Improving Sentence-Level MT Evaluation", Division of Engineering and Applied Sciences Harvard University 33 Oxford St. Cambridge, MA 02138, USA, 75-84.
3. Andrei Popescu-Belis (2002): "An experiment in comparative evaluation: humans vs. computers", ISSCO/TIM/ETI, University of Geneva, 55-64.
4. Andrew FINCH, Eiichiro SUMITA, Yasuhiro AKIBA (2004): "How Does Automatic Machine Translation Evaluation Correlate With Human Scoring as the Number of Reference Translations Increases?" ATR Spoken Language Translation Research Laboratories, 2-2-2 Hikaridai "Keihanna Science City" Kyoto, 619-0288, Japan, 2019-2022.
5. Bohan Niamh, Breidt, Volk Martin, (2000): "Evaluating Translation Quality as Input to Product Development", 2nd International Conference on Language Resources and Evaluation, Athens.
6. Coughlin, D. (2003) "Correlating Automated and Human Assessments of Machine Translation Quality" in *MT Summit IX, New Orleans, USA*, 23–27.
7. Deborah Coughlin, (2003): "Correlating Automated and Human Assessments of Machine Translation Quality", In Proceedings of MT Summit IX. New Orleans, 63-70.
8. Donaway, R.L., Drummey, K.W., and Mather, L.A., (2000): "A Comparison of Rankings Produced by Summarization Evaluation Measures", In Proceedings of the Workshop on Automatic Summarization, 69-78.
9. Feifan Liu, Yang Liu (2008): "Correlation between ROUGE and Human Evaluation of Extractive Meeting Summaries", the University of Texas at Dallas Richardson, TX 75080, USA, 201-208.
10. George Doddington (2002): "Automatic Evaluation of Machine Translation Quality Using N-gram Co-Occurrence Statistics", In Proceedings of the Second Conference on Human Language Technology (HLT-2002). San Diego, CA. 128-132.
11. http://en.wikipedia.org/wiki/Evaluation_of_machine_translation
12. http://en.wikipedia.org/wiki/History_of_machine_translation
13. Jesus´ Angel Gimenez Linare (2008): "Empirical Machine Translation and its Evaluation", Artificial Department the Languages Systems Informatics University, 27-38.
14. Paula Estrella, Andrei Popescu-Belis , Maghi King (2007): "A New Method for the Study of Correlations between MT Evaluation Metrics", ISSCO/TIM/ETI University of Geneva 40, bd. du Pont-d'Arve 1211 Geneva, Switzerland, 35-43.
15. Philipp Koehn (2004): "Statistical Significance Tests for Machine Translation Evaluation" Computer Science and Artificial Intelligence Laboratory Massachusetts Institute of Technology, The Stata Center, 32 Vassar Street, Cambridge, MA 02139.
16. Philipp Koehn (2004): "Statistical Significance Tests for Machine Translation Evaluation" Computer Science and Artificial Intelligence Laboratory Massachusetts Institute of Technology, The Stata Center, 32 Vassar Street, Cambridge, MA 02139.
17. Rao, Durgesh (2001): "Machine Translation in India: A Brief Survey", National Centre for Software Technology Gulmohar Road 9, Juhu, Mumbai 400049, India, 21-23.
18. Sanjay Kumar Dwivedi, Pramod Premdas Sukhadeve (2010): "Machine Translation System in Indian Perspectives" Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Lucknow, India, J. Computer Sci., 6 (10): 1111-1116.

19. Sanjay Kumar Dwivedi, Pramod Premdas Sukhadeve (2010): "Machine Translation System in Indian Perspectives", Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Lucknow, India.
20. Tomer Neeraj and Sinha Deepa, "Evaluating Machine Translation Evaluation's BLEU Metric for English to Hindi Language Machine Translation", in The International Journal of Computer Science & Application, Vol-01-NO-06-Aug-2012, 48-58.
21. Tomer Neeraj and Sinha Deepa, "Evaluating Machine Translation Evaluation's modified BLEU Metric for English to Indian Language Machine Translation", in International Journal of Emerging Sciences (IJES).
22. Tomer Neeraj and Sinha Deepa, "Evaluation of Modified-BLEU Metric for English to Hindi Language Using ManTra Machine Translation Engine", in International Journal of Advanced Research in Electronics & Communication Engineering (*IJARECE*), Volume 1, Issue 4, October 2012, 103-108.
23. Tomer Neeraj and Sinha Deepa: "Evaluating Machine Translation Evaluation's NIST Metric for English to Hindi Language Machine Translation", paper accepted in The International Journal of Multidisciplinary Academy IJMRA for November-2012 issue.
24. Tomer Neeraj, "Evaluating Machine Translation (MT) Evaluation Metrics for English to Indian Language Machine Translation", Ph.D. Thesis 2012, Banasthali University, Banasthali.
25. Tomer Neeraj, Sinha Deepa and Piyush Kant Rai, "Evaluating BLEU Metric for English to Hindi Language Using ManTra Machine Translation Engine", is communicated in International Journal of Advanced Research in Computer Science-IJARCS.
26. Tomer Neeraj, Sinha Deepa and Piyush Kant Rai, "Evaluating Machine Translation Evaluation's F-Measure Metric for English to Hindi Language Machine Translation", in International Journal of Academy Research Computer Engineering and Technology-IJARCET, volume1, Issue 7, September 2012, 151-156.
27. Vannatta Rachel: "Statistics in Education Course Packet", EDFI 641. Online available http://personal.bgsu.edu/~rvanna/packetspring 09.pdf.

Author 1

Neeraj Tomer
tneeraj12@rediffmail.com
9460762117

Area of Interest:

- Machine Translation and Indian Language Technology
- Theoretical Computer Science and related technologies

Academic Qualification:

Ph.D (thesis submitted) in Computer Science, Banasthali University, Banasthali.
MCA, Maharishi Dayanand University, Rohtak 2005.
Master of Economics, Kurukshetra University Kurukshetra 1999.
Bachelor of Economics, Kurukshetra University Kurukshetra 1997.

Employment History:

Post graduate and graduate teaching at Mahatma Gandhi Institute of Applied Sciences, Jaipur as a lecturer from July 2003 to August 2006. As a Research Associate at Banasthali University Banasthali in 2007. As a lecturer at LCRT College of Education Panipat from August 2007 to July 2010. As an Assistant Professor at SINE International Institute of Technology, Jaipur from August 2010 to March 2012.

- Papers Published : 2
- In Press : 2
- Communicated : 3

Seminar and Conferences Attended: 5

Research Guidance:

Guided 3 students for their dissertation work at PG (M.Sc) level.

Future Plans: To grow academically

Author 2

Deepa Sinha
Associate Professor
Department of Mathematics
South Asian University
Akbar Bhawan
Chanakyapuri, New Delhi 110021 (India)
Cell No: 08744022273
deepasinha2001@gmail.com

Research Area: **Graph Structures**

Academic Qualification:

M.Sc., Ph. D. (University of Delhi), CSIR-NET (twice)

Future Plans: To grow academically

Achievements: CSIR_NET (qualified Twice)

Publications:

(a) Books: one
(b) Research Papers: 27

Conference/workshop/symposium attended: 39

Invited talks delivered: Ten

Papers presented: 23

Work experience: 16 years

Served several Institutions for graduate and post graduate courses, particularly Khalsa Girls Degree College (1996-1999), Delhi College of Engineering (1999-2004), Banasthali University (2004-2012).
Seven students got awarded their M. Phil. Degree under her supervision.
Three students got awarded their Ph.D. in the year 2011-2012.
Have refereed several research papers for National and international Journal of high impact factor like Discrete Applied Mathematics, Graphs and Combinatorics, International Journal of Physical Sciences etc.
Sessions chaired in the National/ International conferences: four

Experimental Determination of Contact Velocity

Dulal Goldar
Distinguished Professor, School of Engineering & Technology, Sharda University, Greater Noida.
Former Principal, Delhi College of Engineering
E-mail: dulalgoldar@yahoo.co.in

Abstract:
The study of transverse impact of beams is important from designer's point of view. The author has done experiments using dynamic photoelasticity as well as strain gauge on model beams made of different materials. For the measurement of contact velocity he also developed a miniature transducer [1]. In all experiments low velocity (> 2 m/s) was generated by free fall of a striker and keeping constant beam/ striker weight ratio (2.675), height of fall of the striker (176.4 mm) and beam dimensions. Two sets of experiments were performed for the measurement of contact velocity namely, for dynamic photoelasticity and dynamic force and strain measurements.

Key Words
Contact-velocity, Impact, Simply supported beam, Dynamic Photoelasticity, Electrical resistance stain gauge, Contact-force Transducer.

Introduction:
The evaluation of contact velocity through experiment is always a challenge for experimentalist. At many times difficulties are faced. In a view to help the research worker in the same field of dynamic recording of data detailed experimental method is presented. A simply supported beam with equal over-hang is common structural element for bridge girder, chassis of railway wagon and automobiles. Such structures are subjected to a transverse impact in service life.
Using Fastax framing camera isochromatic fringe photographs were recorded. Urethane rubber beam (PSM-4) (253 mm x 24.3mm x 12.5mm thick) was used. Locations of the supports were changed to generate three different simply supported beams namely, 90mm, 120mm, 150mm beam spans. A freely falling mass (Fig.1a and 1b) inducing a low velocity (< 2m/s) impact at central- and quarter- span locations. In all beam striker weight ratio [2], [3] was kept constant i.e.2.675.

Experimental Set- up:
The first experimental set-up (Fig.1.) essential equipments are listed in Fig.1. For the sake brevity the same are not repeated.
For recording dynamic isochromatic fringe pattern author employed Fastax 16mm film framing camera (operating 12,000 f.p.s). The plane polarizer and quarter wave plates were placed in such a manner that full- light-field arrangement was created for recording isochromatic fringe photograph. This helped to record movement of the striker on the film plane through a slot cut diametrically at the end of guide pipe (Fig.1 and Fig.3).

Method of Synchronization:
1. A schematic diagram showing the method of synchronization used in fringe photography is indicated in Fig.2. Before recording a particular fringe photograph, the following operations were conducted in sequence.
2. The time of fall of the striker (Fig.1a and 1b) was measured by "starting" a timer circuit by operation of the electromagnet operating switch and "stopping" the timing circuit by a metal foil short- circuiting. The time of fall was used in setting "Event marking" time of the "Goose- control Unit" (Fig.2).
3. Depending upon the framing speed desired and total length of the film used, the "Voltage Control" and "Camera time" of the "Goose control unit" were then set [4].
4. The Relay (Fig.2) was at "Normal" position by which the electromagnet was energized.
5. All the laboratory lights were switched off and Sun- gun toggle switch was on.
6. The "Press-button" switch of the "Time-marker" was pressed.
7. The "Micro-switch" (Fig.2) was operated which set the Relay in "Record" position. The Relay de-energized the electromagnet and simultaneously switching on the "Goose-control unit". The "Goose- control unit" in turn started motors of the framing camera. By the time the striker hit the beam model, the desired camera speed was achieved. The camera was automatically switched off by the "Camera Time" control provided in the "Goose- control unit", and therefore, the recording was completed.
8. The "Time-marker" press-button switch and "Sun- gun Toggle switch" were then switched off.

Photographic Film:
KODAK2498RAR (250ASA) 16 mm roll films were used for fringe photography. KODAK developer, freshly prepared from chemicals, was used for developing these film rolls.
Enlarged prints, both continuous (Fig.3a, 3b) and discrete (selected frames) (Fig.4), were prepared from the film negatives to study the history of fringe formation and to identify precisely fringe orders, respectively.

Experimental Details:

The camera produced sparks 1000 cycles per second on the 16 mm film negative (Fig.3).Number of frames between these white dots was recorded to obtain inter-frame time (I.F.T.).

A full- field photoelastic stress analysis on a rubber- like material (PSM-4) is first conducted by the techniques of dynamic photoelasticity to get an insight into stress distribution for the problem at hand and thereby identify the points of special interest from the designer's point of view; A freely falling mass (striker) (Fig.1a & 1b) was employed to generate transverse impact on beams. A vertical slot was created at the end of guide pipe through which the striker was allowed free fall. The movement of the striker was clearly visible on the photo-negative. Details of such movement stage-wise are presented the Fig.3a and 3b. Schematically five stages are explained here in Fig.3a and Fig.3b.

The number of frames were counted until the striker touched the top surface of the beam and the same was identified from the photo-negative. The data for movement of the striker was measured on a motion analyzer (Fig.5). The same was fed to a computer to fit a straight line. The slope the straight line was measured to get the contact velocity. Three cases i.e. different beam span were studied, namely 90 mm, 120 mm and 150 mm for central as well as non-central transverse impact studies. In all the cases the contact velocity varied (Table-1, Table-3) but always it was less than 2 m/s.

Results of the contact velocity from this experiment are shown in Table-1. There were some variations in the values of contact velocity of the striker. For the sake of brevity isochromatic fringe pattern (continuous print) for 150 mm beam span subjected to central and quarter- span impact loading is shown in Fig.3a and 3b and discrete enlarged isochromatic fringe photograph for 120 mm beam span under central impact loading is shown in Fig.4. From this photograph isochromatic fringe definition is very clear.

For the second set of experiments (Fig.6.) essential equipments are listed in Fig.6. For the sake brevity the same are not repeated.

A point- wise study employing electrical resistance strain gauges is then conducted at the point of special interest on model beams made of materials like ploymethyl-metha-acrylate (PMMA). In the second set of experiment author used a guide pipe and inserted hole on diametrically opposite side of the guide pipe. A dc powered light source was placed on one side of the guide pipe (Fig.6) and on another side a photo-transistor (detector) was placed across the two holes made diametrically. As the striker was allowed to fall freely through the guide pipe the light path was intercepted by the striker. The photo- transistor (detector) will provide a signal to start the storage oscilloscope. It gave a horizontal sweep till the striker hits the top of the contact force transducer placed over the beam. Time was measured from the oscilloscope horizontal trace. The distance of the position of hole was known and the contact velocity was measured. The contact velocity (Table-2 and 3) was always less than 2 m/s.

Conclusion:

The experimental results are presented in Table-1, Table-2 and Table-3. The I.F.T and contact velocity were different for different experiments on urethane rubber beam as well as other experiments of beams made of other materials. But contact velocity was always remained less than 2 m/s. Contact velocity as shown in Table-3 for central and quarter-span impact loading measured from fringe photographs and oscilloscope traces are almost same.

References:

[1] Goldar,D.,et.al."Development and calibration of a dynamic contact- force transducer", *Proc. 40th. Anniversary meeting, spring meeting, Society for Experimental Stress Analysis, Cleveland, Ohio, USA and subsequently pub. J.Exp.Mech., September*, 1985,pp.187-190.

[2] Goldar, Dulal, "Photoelastic Studies of Transversely Impacted Simply Supported Beams", *Ph.D. Thesis*, August 1981, *Panjab University, Chandigarh.*

[3] Goldar, D., Verma S.P. and Paldas, M "Photoelastic studies of transversely impacted beams with over-hang", *Proc. Symposium on 'War - Head Technology' Organized by Terminal Ballistic research Laboratory (TBRL), Chandigarh*, March 1983.

[4] Operators Manual, *Fastax High Speed Cameras, Category-I through VI, Wollensak division, Ravere Camera Co. Rochester,NY,USA.*

Table-1 Inter-frame-time & Contact-velocity from Photoelasticity

Over-hang Ratio	Span (mm)	Striker weight (gm)	Height of fall (mm)	guide pipe & beam GAP (mm)	Central- Impact		Quarter-span Impact	
					I.F.T. (μs)	Contact velocity (m/s)	I.F.T. (μs)	Contact velocity (m/s)
0.322	90	10.52	176.4	13.9	94.86	1.446	104.80	1.422
0.263	120	1.02	176.4	10.0	84.11	1.601	105.96	1.700
0.204	150	17.53	176.4	14.9	96.78	1.472	96.02	1.163

Table-2 Contact-velocity from from Oscilloscope traces

Material	Striker mass	Beam weight/striker wt.	Contact velocity (m/s)
Aluminum	41.5	2.468	1.827
Perspex(PMMA)	41.5	1.079	1.675
Perspex-Al-Perspex	41.5	1.540	1.724

Table-3 Contact-velocity from Oscilloscope traces and Fringe photography
For urethane rubber beam model

Cases	Striker mass (gm)	Beam span (mm)	Over-hang Ratio	Span/Depth Ratio	Contact Velocity(m/s)	
					Osc.traces	Fringe Photo
a.	10.52	90	0.322	3.711	1.450	1.466
b.	14.02	120	0.263	4.948	1.658	1.601
c.	17.53	150	0.204	6.186	1.460	1.472
p.	10.52	90	0.322	3.711	1.450	1.422
q.	14.02	120	0.263	4.948	1.718	1.700
r.	17.53	150	0.204	6.186	1.460	1.163

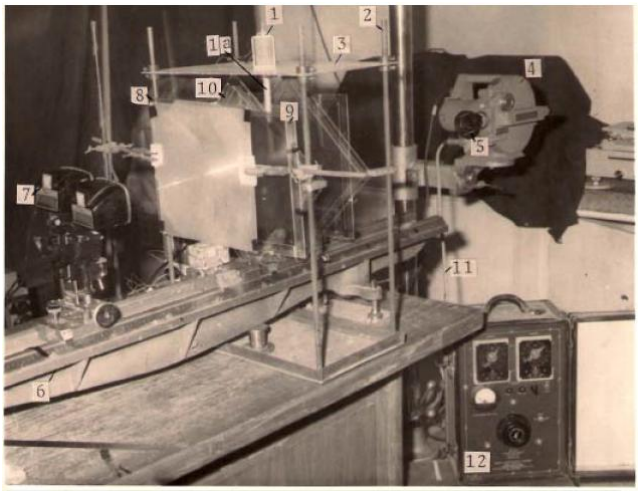


Fig.1 Photograph of Experimental Arrangement for Fringe Photography
Legend:1Electromagnet, 1aGuide pipe, 2Vertical m.s rod suppo, 3.Aluminum plate, 4.Fastax, 5Monochromator, 6 Optical bench, 7Light source, 8 Fresnel lens, 9Plane Polaroid, 10Quarter wave plate, 11Signal from time marker, 12Goose control unit

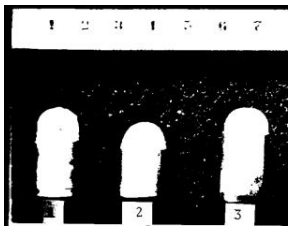


Fig.1 (a) Photograph of 1,2 & 3 Strikers with hemispherical Araldite tip of weight 10.52gm, 14.02gm and 17.53gm)

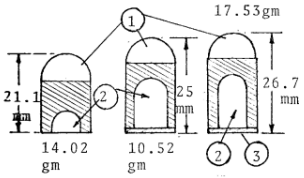
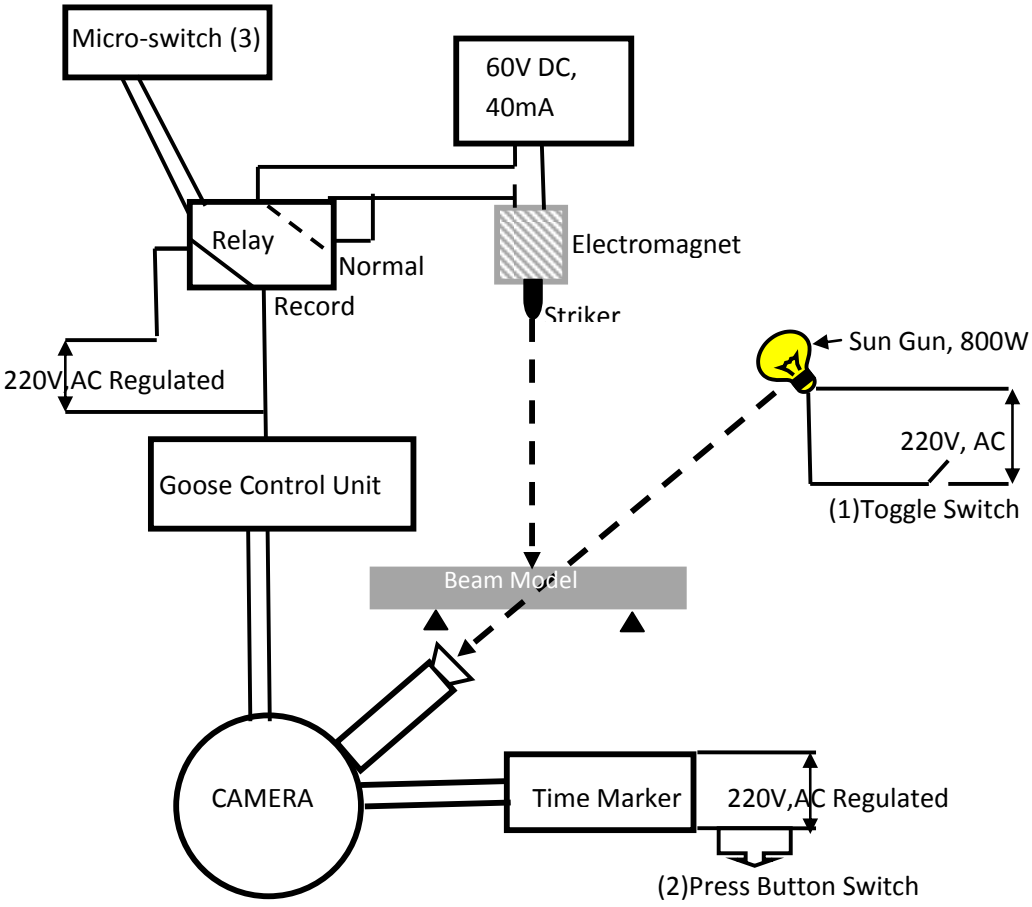


Fig.1b Details of Striker with hemispherical Araldite tip of weight 10.52 gm, 14.02 gm and 17.53 gm. Cavity for adjusting weight of striker mild steel disc for holding properly with

Fig.2 Schematic Diagram Showing Method of synchronization



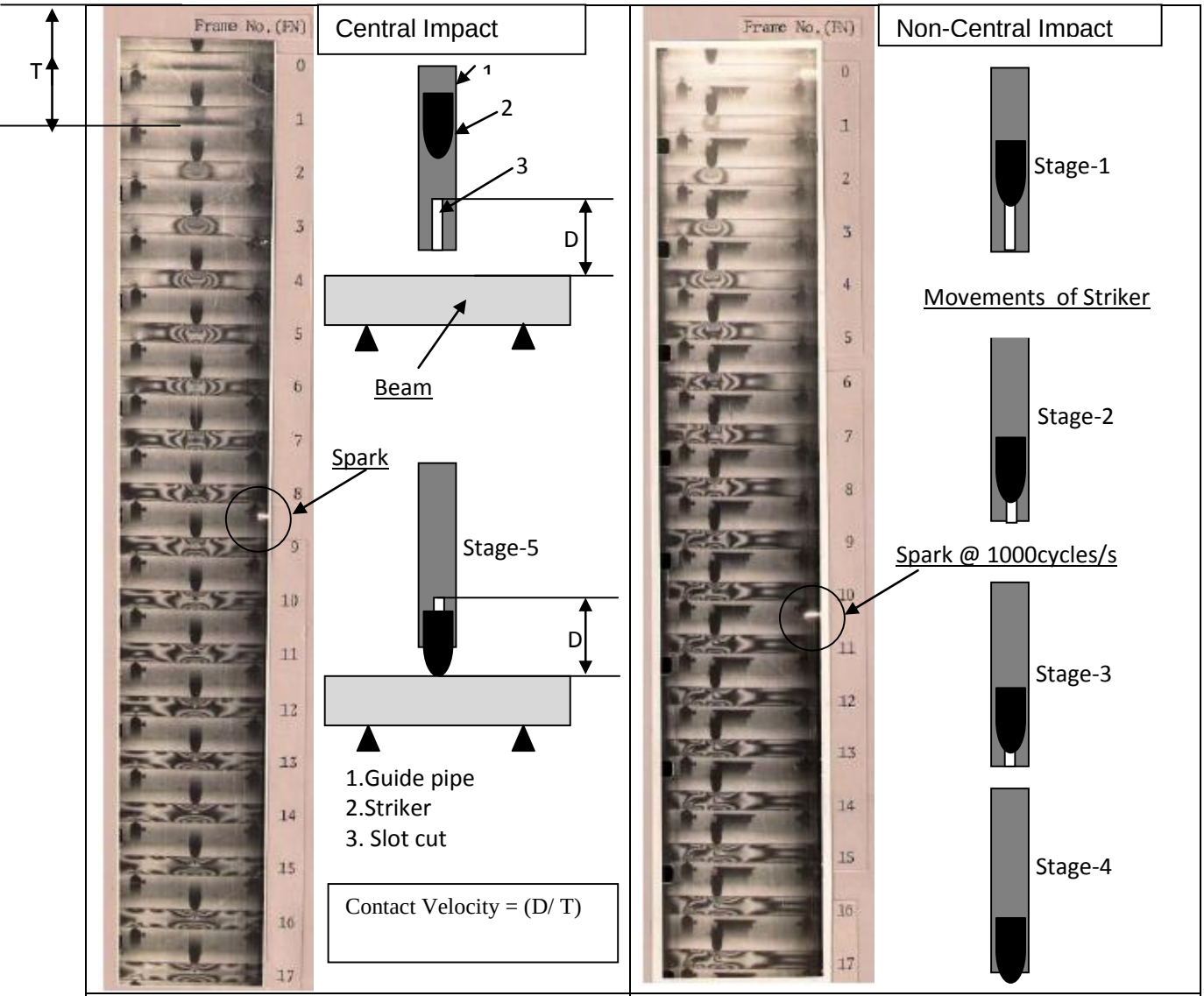


FIG.3(a) Isochromatic Fringe Photography
Span 150mm, over-hang ratio=0.204, IFT=96.78μs

FIG.3(b) Isochromatic Fringe Photography
Span 150mm, over-hang ratio=0.204, IFT=96.02μs

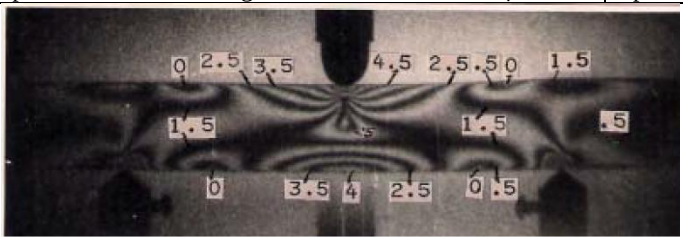


Fig.4 Enlarged Isochromatic fringe photograph of 120 mm - Beam Span Central Impact, Frame No. 30, Time = 2523μs

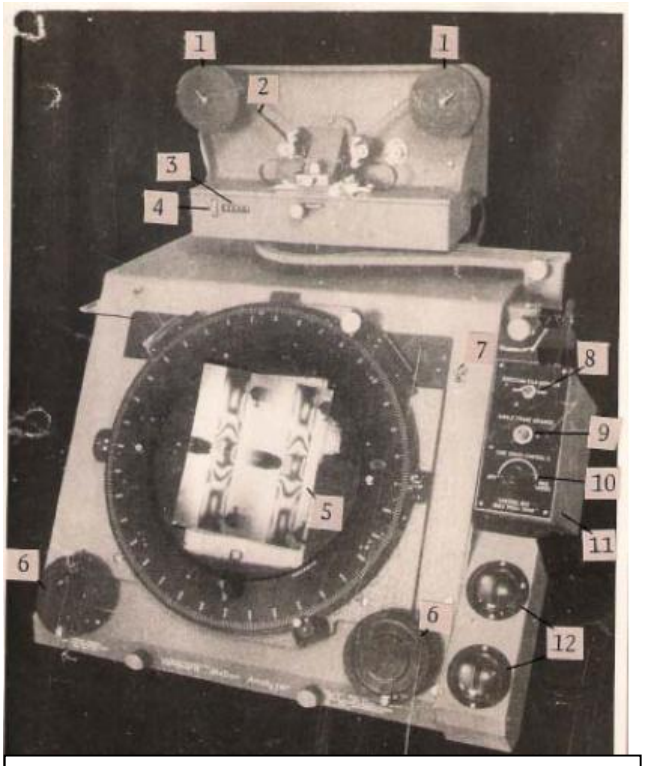


Fig.5 Motion Analyser

Legend: 1. Film Holdaer(Spool), 2.Film(16mm), 3.Digits showing Frame Number, 4.Zero Adjusting Knob, 5.Screen(16.5xmagnification), 6.Knob for Hor. (X)/ Vert. (Y) movements of cursor, 7.On-Off Switch, 8.Direction for Film, 9.Single Frame Advance, 10.Cine Speed Control, 11.Control Box, 12.Digital Read out for (X Readings/ (Y) Readings

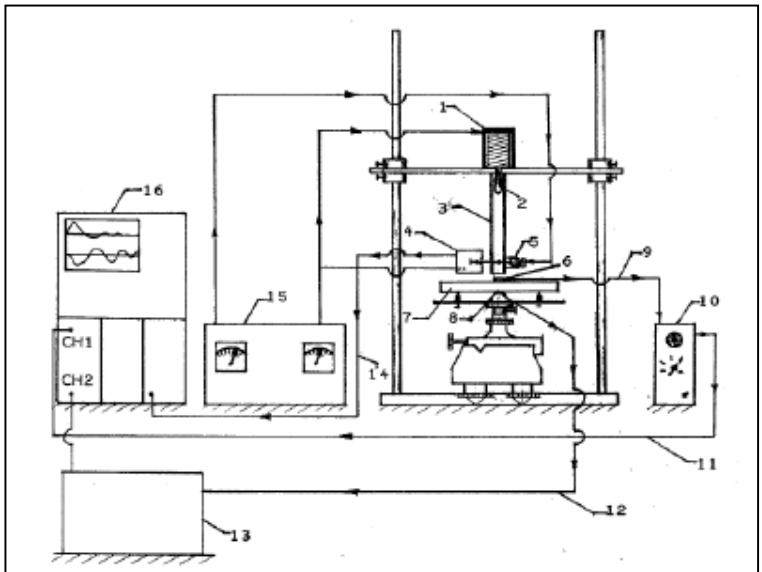


Fig.6 Experimental Arrangement For Measurement of Contact, Strain & Contact Velocity

Legend: 1.Electromagnet, 2. Striker, 3. Guide Pipe, 4. Phototransistor
5. Light source (10V DC), 6. Contact force transducer, 7. Beam Model, 8. Strain Gauge, 9. Signal from Force transducer
10.Charge amplifier (KISTLER), 11. Signal from charge amplifier
12. Signal from strain Gauge, 13. Bridge Balancing & Amplifying Unit, 14. Signal from Phototransistor to external trigger
15. Variable DC power supply, 16. Storage oscilloscope

Fig.7a. Oscilloscope Trace

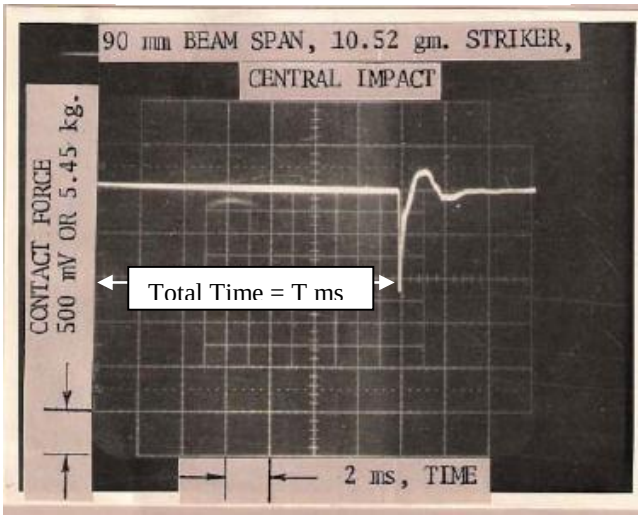


Fig.7b Details of Measurement

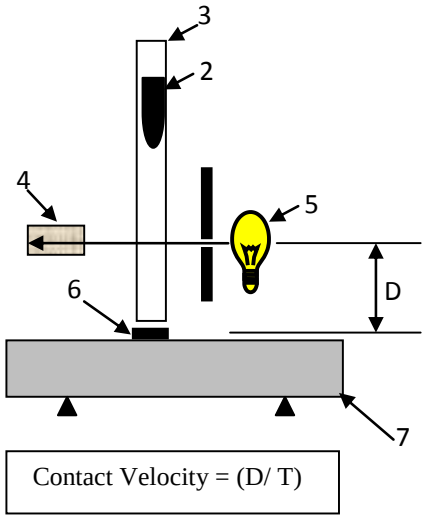


Fig. 7 Typical Oscilloscope Trace Showing Contact-Force History for Urethane Rubber Beam Model

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 1 Issue 8, October - 2012

Improvement In Power System State Estimation By Use Of Phasor Measurement Unit

Jitender Kumar*[1], J.N.Rai[2], Vipin[3], B.B. Arora[3] and C.K.Singh[1]

[1]*Electrical Engineering Deptt, G L Bajaj Instt of Tech & Management, Gr. Noida, India*
[2]*Electrical Engineering Department, Delhi Technological University, Delhi, India*
[3]*Mechanical Engineering Department, Delhi Technological University, Delhi, India*

Abstract

This paper mainly focuses on a placement algorithm of phasor measurement unit which enhance the accuracy of the state estimation. The algorithm developed will ascertain a list of buses with low quench of estimator accuracy. The optimal placement of phasor measurement unit is determined with respect to their characteristic of parameters measured by using phasor measurement unit. The test results of the above are tested on IEEE 6-bus system, IEEE 9-bus system and IEEE 14-bus system.

Keywords-- **phasor measurement unit, state estimation, weighted least square.**

1. Introduction

The concept of state estimation is originally proposed by Schweppe in 1970s. Over the last 3 decades of power system, the state estimation will help in resolving the parameter of power system. The State estimation is conventionally explained by the weighted least square algorithm (WLS) [1] with conventional measurements of system voltage magnitude and system current.

The Recently developed technique of phasor measurement unit will able to measure the system voltage and system current in the power system. It produces a great impact on the state estimation parameter measurement. In reference [8], a detailed review on development of the phasor measurement unit made by Phadke. The phasor measurement unit at substation will allows us for direct measurement of the state of the network. It makes the whole state estimator as a linear estimator. However, it may be reasonable in replacement of all the existing measurements system with phasor measurements in the power system due to accuracy and quickly analyzing the behavior of system parameter.

The main objective is to optimally situate a phasor measurement unit to obtain maximum benefit in regard to state estimation. Several methods [2] are proposed in relation of PMU placements as annealing method, graph theoretic procedures, depth first, tabu search method and genetic algorithms. They are proposed in determination of a minimum set of phasor measurement unit to make the network more observable. There is also several meter assignment methods proposed in literature in determination of an optimal set of measurements. The criterion of determining the minimum number of the measurement matrix proposed in an optimal meter placement [11] method.

The authors [14] concentrate on the accuracy, redundancy and cost requirements while determine the placement of measurement devices. In an incremental meter placement [9], the State Estimation solutions will improve their accuracy and robustness. The criterion of variances of state estimation errors is proposed in incremental meter placement.

In this paper, a placement algorithm for phasor measurement unit is proposed. Its objective is to improve the accuracy and robustness of the solution of state estimation when an inadequate number of PMUs can be added to the power system. The algorithm will determine the phasor measurement unit placement by using the main three aspects as: system topological structure, accuracy and redundancy. The algorithm has been tested on the IEEE 6-bus test system [7], IEEE 9-bus test system [3] and IEEE 14-bus test system [13].

Phasor Measurements Unit

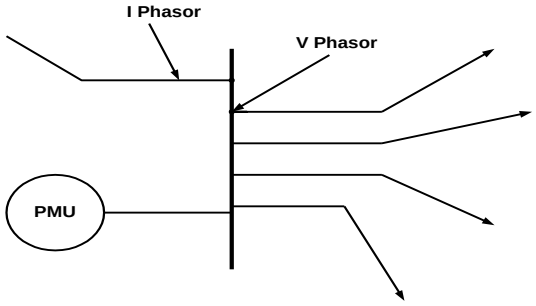


Fig. 1: Phasor Measurements Unit

For the practical purposes the Fig. 2 will shows a 4-bus system, which has single PMU at bus 1. It

has one voltage phasor measurement and three current phasor measurements [12], namely $V_1 \angle \theta_1$, $I_1 \angle \delta_1$, $I_2 \angle \delta_2$ and $I_3 \angle \delta_3$.

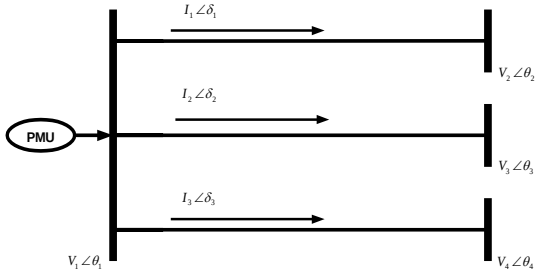


Fig 2: Single PMU Measurement Model

The investigators of Virginia Polytechnic Institute [8] will develop the phasor measurement unit in around 1970s. They state that the PMU has capability of providing a synchronized real time measurement of current phasors and voltage phasors. The Synchronization within the buses is achieved at the same-time sampling of current and voltage waveforms using by timing signals from the Global Positioning System [10]. A PMU placed on a bus makes that the bus and all its fellow buses are observable. This compose the incremental placement of PMU are different from standard measurements placement [4].

A brief review of state estimation in power system is presented in section 2. The proposed algorithm in relation to placement of phasor measurement unit is presented in Section 3. The test results are carried out on the IEEE 6-bus, IEEE 9-bus and IEEE 14-bus test systems are given in Section 4. Finally some concluding notes are presented in Section 5.

2. Power System State Estimation

The WLS state estimation minimizes the weighted sum of squares of the residuals. Consider the set of measurements given by the vector *z:*

$$z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ . \\ . \\ z_m \end{bmatrix} = \begin{bmatrix} h_1(x_1, x_2, x_3, ...x_n) \\ h_2(x_1, x_2, x_3, ...x_n) \\ h_3(x_1, x_2, x_3, ...x_n) \\ . \\ . \\ h_m(x_1, x_2, x_3, ...x_n) \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ . \\ . \\ e_m \end{bmatrix} = h(x) + e \quad (1)$$

Where *z* is the measurement vector (*m*-vector), *x* is the true state vector (*n*-vector, *n*<*m*)

$$h^T = [h_1(x), h_2(x), h_3(x), ...h_m(x)] \quad (2)$$

h_i *(x)* is the nonlinear function relating measurement *i* to the state vector *x*

x^T = [x_1, x_2, x_3..... x_n] is system state vector

e^T = [e_1, e_2, e_3…….e_m] is the vector of measurement errors.

The Weighted Least Square estimator will minimize the following objective function:

$$J(x) = \sum_{i=1}^{m} \frac{(z_i - h_i(x))^2}{R_{ii}} = [z - h(x)]^T R^{-1} [z - h(x)] \quad (3)$$

The non-linear function $g(x)$ can be expanded into its Taylor series around the state vector x^k neglecting the higher order terms.

$$g(x) = g(x^k) + G(x^k)(x - x^k) + ... = 0 \quad (4)$$

An iterative solution scheme known as the Gauss-Newton method is used to solve above equation:

$$x^{k+1} = x^k - [G(x^k)]^{-1} . g\left(x^k\right) \quad (5)$$

where, *k* is the iteration index and x^k is the solution vector at iteration *k* . *G*(*x*) is called the gain matrix, and expressed by:

$$G(x) = \frac{\partial g(x^k)}{\partial x} = H^T(x^k) R^{-1} H(x^k) \quad (6)$$

$$g(x^k) = -H^T(x^k) R^{-1} [z - h(x^k)] \quad (7)$$

Generally, the gain matrix is quite sparse and decomposed into its triangular factors. At each iteration *k*, the following sparse linear sets of equations are solved using forward/backward substitutions, where $\Delta x^{k+1} = x^{k+1} - x^k$:

$$[G(x^k)]\Delta x^{k+1} = H^T(x^k)R^{-1}[z-h(x^k)] = H^T(x^k).R^{-1}\Delta z^k \quad (8)$$

These iterations are going on until the maximum variable difference satisfies the condition, ' *Max* $|\Delta x^k| < \varepsilon$ '. A detailed flow-chart of this algorithm is shown in next section.

3. The Proposed PMU Placement Algorithm

The objective of the proposed algorithm in relation to placement of PMU is to reduce the variances of the Weighted Least Square State Estimation error. Also help in increases the local redundancy at the same time. While a PMU placed on a specific bus makes that bus and its entire neighbor buses are observable, buses are formed a group by using that single PMU placement [5]. The weakest bus group is determined by the local redundancy ranking and Weighted Least Square State Estimation error variance ranking. The local redundancy rankings are determined by the method described in [1]. The ranking of accuracy are determined by the diagonal elements of the Weighted Least Square State Estimation covariance matrix.

The WLS with PMU algorithm is illustrate as follows:

1. Start iterations, set the iteration index k = 0.
2. Initialize the state vector x^k, typically as a flat start.
3. Calculate the gain matrix, $G(x^k)$.
4. Calculate the right hand side $t^k = H^T(x^k)R^{-1}[z-h(x^k)]$
5. Decompose $G(x^k)$ and solve for Δx^k
6. Test for convergence, *Max* $|\Delta x^k| < \varepsilon$?
7. If no, update $x^{k+1} = x^k + \Delta x^k$, k = k + 1 and go to step 3. Else, stop.

The above algorithm essentially involves the following computations in each iteration, k;

1. Calculation of the RHS parameter $t^k = H^T(x^k)R^{-1}[z-h(x^k)]$
   (a) Calculating the measurement function, h (x^k).
   (b) Building the measurement Jacobian, H (x^k).
2. Calculation of G (x^k) and solve for Δx^k.
   (a) Building the gain matrix, G (x^k).
   (b) Decomposing G (x^k) into its Cholesky factors.
   (c) Performing the forward/back substitutions to solve for Δx^{k+1}.

4. Simulation Results

In this section, the proposed algorithm in relation to placement of PMU in the previous section has been tested on the IEEE 6-bus test system [7], IEEE 9-bus test system [3] and IEEE 14-bus test system [13]. In respect of the system accuracy and reliability, PMU can deliver more precise measurement data [6]. Several cases to be tested with PMUs added to the conventional measurement set. The simulations and analysis of different cases are as shown in Table 1 are done with several IEEE bus systems in the next section.

TABLE 1
DIFFERENT CASES PMU ADDITION IN IEEE SYSTEM

Cases	Measurements
1	Conventional with No PMUs
P	Only PMUs

For investigate the system accuracy with or without PMU on system variables, some cases are tested with the help of MATLAB simulink software. The testing parameters are available on conventional method with or without PMU.

TABLE 2
PMU LOCATIONS FOR IEEE SYSTEM

Type of System	PMU locations at Bus
IEEE 6 System	Bus 2
IEEE 9 System	Bus 2
IEEE 14 System	Bus 2

The circuit diagram will be shown as in Fig.3 for IEEE 6.

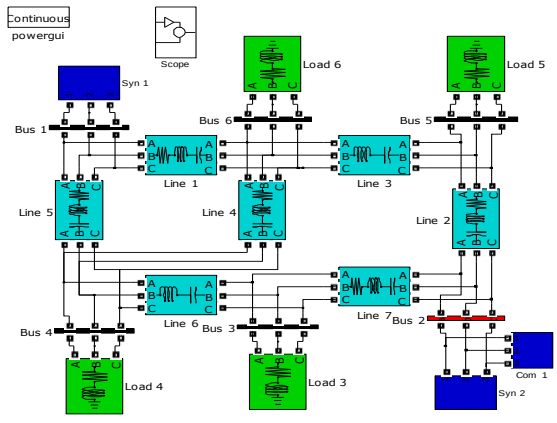


Fig. 3: IEEE 6 Bus System

Similarly the placement of PMU to be done on IEEE 9-bus test system and IEEE 14-bus test system.

In this segment, IEEE bus systems as IEEE 6 bus system, IEEE 9 bus system and IEEE 14-bus

test system are tested with their respective cases to find out the consequences of the PMUs to the precision of the estimated variables.

The settings for error standard deviations for measurements are shown in Table 3. A PMU has much smaller error deviations than conventional measurements as 0.0001.

TABLE 3
MEASUREMENT S.D. FOR THE TEST

Real Power (Max)	Real Power (Min)	Reactive Power (Max)	Reactive Power (Min)
0.001	0.0001	0.0001	0.0001

The parameters measured are Real Power and Reactive Power (flow & injected) measurements. The variation of parameters with or without PMU easily reflected in the fig. 4 – 15 as below:

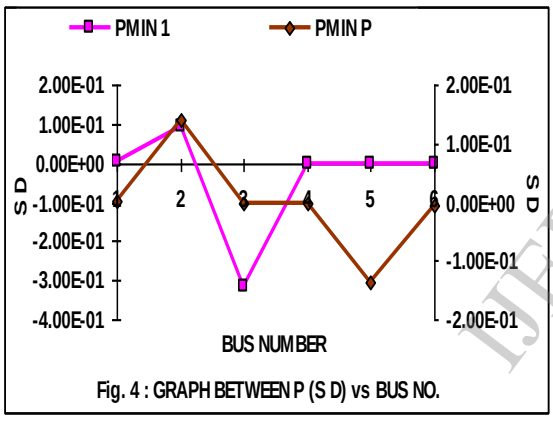


Fig. 4 : GRAPH BETWEEN P (S D) vs BUS NO.

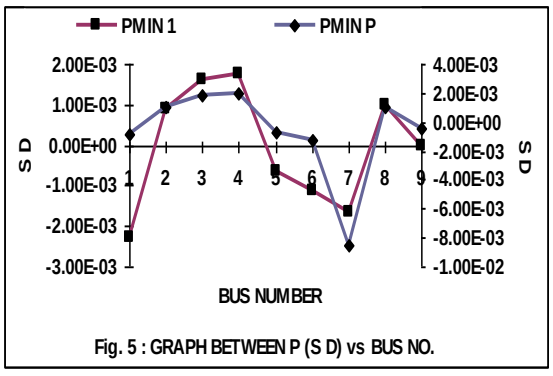


Fig. 5 : GRAPH BETWEEN P (S D) vs BUS NO.

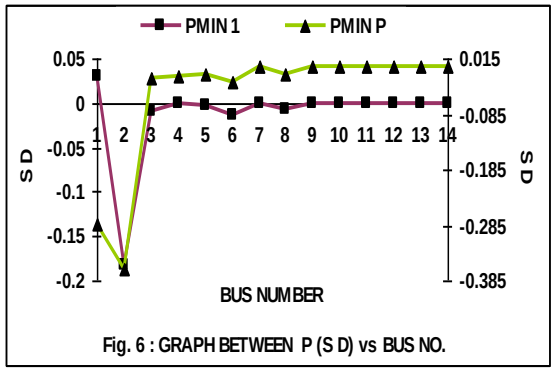


Fig. 6 : GRAPH BETWEEN P (S D) vs BUS NO.

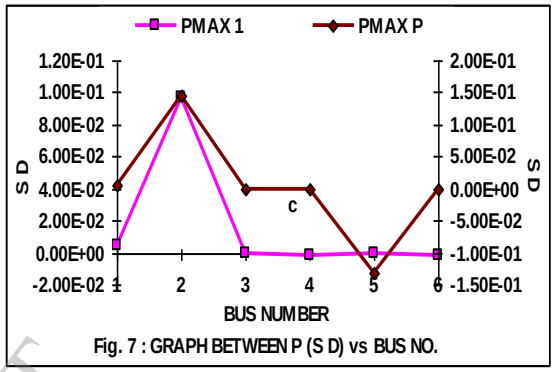


Fig. 7 : GRAPH BETWEEN P (S D) vs BUS NO.

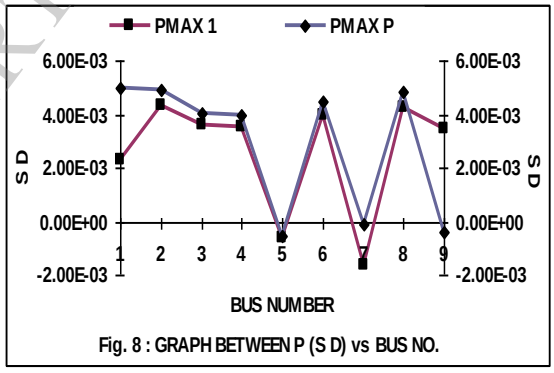


Fig. 8 : GRAPH BETWEEN P (S D) vs BUS NO.

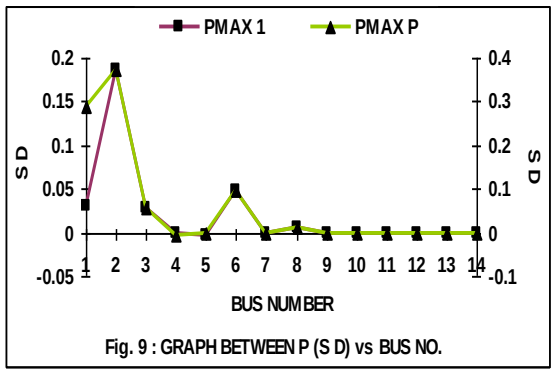


Fig. 9 : GRAPH BETWEEN P (S D) vs BUS NO.

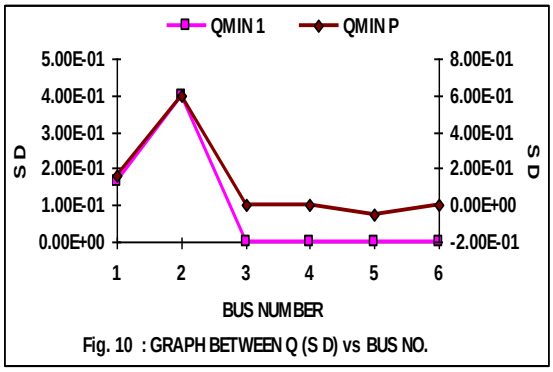


Fig. 10 : GRAPH BETWEEN Q (S D) vs BUS NO.

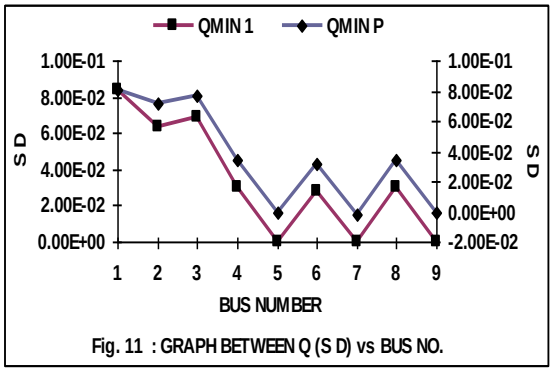


Fig. 11 : GRAPH BETWEEN Q (S D) vs BUS NO.

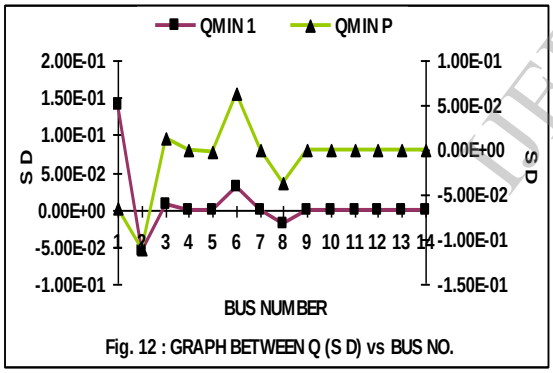


Fig. 12 : GRAPH BETWEEN Q (S D) vs BUS NO.

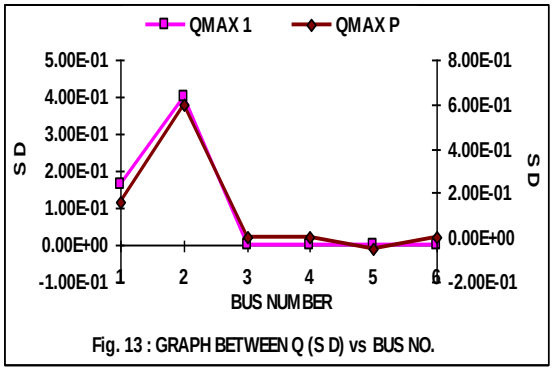


Fig. 13 : GRAPH BETWEEN Q (S D) vs BUS NO.

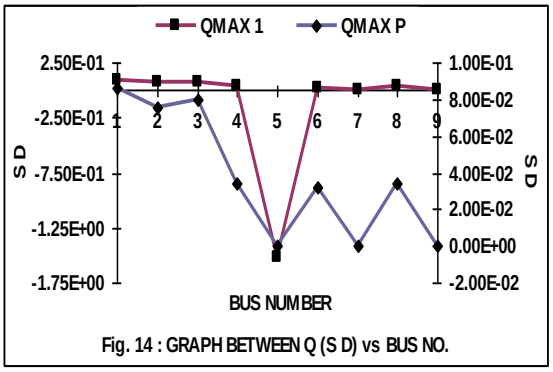


Fig. 14 : GRAPH BETWEEN Q (S D) vs BUS NO.

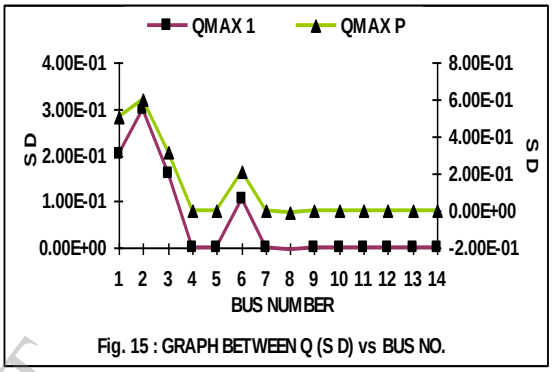


Fig. 15 : GRAPH BETWEEN Q (S D) vs BUS NO.

5. Conclusions

An algorithm in relation to PMU placement for power system state estimation has been presented. The algorithm establishes the optimal placement of PMU to exterminate the critical measurement. While at the same time it develops the accuracy of Weighted Least Square State Estimation and its measurement with their effective redundancy level. The algorithm determines the optimal placement of the PMU based on criterion which takes into account critical measurement and PMU characteristics. The test results on IEEE 6-bus system, IEEE 9-bus system and IEEE 14-bus test system shows that the addition of PMU measurement will improves the performance of the power system state estimation significantly at real time configurations.

6. References

[1] Abur and Exposito A. G., Power System State Estimation, Theory and Implementation, MAECEL DEKKER, 2005, pp. 9-27.

[2] Dongjie Xu; Renmu He; Peng Wang; Tao Xu, "Comparison of several PMU placement algorithms for state estimation", 2004.Eighth IEEE International Conference on Developments in Power System Protection, Vol. 1, 5-8 April 2004; pp: 32 - 35 Vol.1

[3] Ebrahimpour R., Abharian E. K., Moussavi S. Z. and Birjandi A. A. M., "Transient Stability

Assessment of a Power System by Mixture of Experts", International Journal of Engineering, (IJE) Vol. 4: Issue (1) March 2011 pp.93–104.

[4] Klump R., Wilson R.E. and Martin K.E., "Visualizing Real-Time Security Threats Using Hybrid SCADA / PMU Measurement Displays", Proceedings of the 38th Hawaii International Conference on System Sciences – 2005, 0-7695-2268-8/05 © 2005 IEEE, pp.1–9

[5] Madtharad C., Premrudeepreechacharn S., Watson N.R. and Saeng-Udom R., "An Optimal Measurement Placement Method for Power System Harmonic State Estimation", IEEE Transactions on Power Delivery, Vol. 20, No. 2, April 2005, 0885-8977 © 2005 IEEE, pp.1514-1521

[6] Madtharad C., Premrudeepreechacham S., Watson N. R. and Saenrak D., "Measurement Placement Method for Power System State Estimation: Part I", 0-7803-7989-6/03 ©2003 IEEE, pp.1629–1632

[7] Meshram S. and Sahu O.P., "Application of ANN in economic generation scheduling in IEEE 6-Bus System", International Journal of Engineering Science and Technology (IJEST). ISSN : 0975-5462 Vol. 3 No. 3, 2011, pp.2461–2466

[8] Phadke A.G., "Synchronized Phasor Measurements – A Historical Overview", 0-7803-7525-4/02 © 2002 IEEE, pp.476–479

[9] Phadke A.G, Thorp J.S., Nuqui R.F. and Zhou M., "Recent Developments in State Estimation with Phasor Measurements", 978-1-4244-3811-2/09 ©2009 IEEE, pp.1-7

[10] Rahman K. A., Mili L., Phadke A., Ree J. D. L. & Liu Y., "Internet Based Wide Area Information Sharing and Its Roles in Power System State Estimation", 0-7803-6672-7/01 © 2001 IEEE, pp.470–475

[11] Rakpenthai, C.; Premrudeepreechacharn, S.; Uatrongjit, S.; Watson, N.R, "Measurement placement for power system state estimation by decomposition technique", 2004. 11th International Conference on Harmonics and Quality of Power, 12-15 Sept. 2004 pp:414 – 418

[12] Real time dynamics monitoring system [Online]. Available: http://www.phasor rtdms.com

[13] Tanti D.K., Singh B., Verma M.K., Mehrotra O. N., "An ANN based approach for optimal placement of DSTATCOM for voltage sag mitigation", International Journal of Engineering Science and Technology (IJEST), ISSN : 0975-5462 Vol. 3 No. 2, Feb 2011 pp.827–835

[14] Zhu Hongbo; Baran, M.E.; Zhu Jinxiang; Garren, K.E, "A meter placement method for state estimation", IEEE Transactions on Power Systems, Volume 10, Issue 3, Aug. 1995 pp :1704 - 1710

International Scholarly Research Network
ISRN Electronics
Volume 2012, Article ID 473257, 7 pages
doi:10.5402/2012/473257

Research Article

MOS Current Mode Logic with Capacitive Coupling

Kirti Gupta,[1] Neeta Pandey,[1] and Maneesha Gupta[2]

[1] *Department of Electronics and Communication Engineering, Delhi Technological University, Delhi-110042, India*
[2] *Department of Electronics and Communication Engineering, Netaji Subhas Institute of Technology, New Delhi-110078, India*

Correspondence should be addressed to Kirti Gupta, kirtigupta22@gmail.com

Received 31 August 2012; Accepted 19 September 2012

Academic Editors: D. Al-Khalili, G. Maruccio, and Z.-M. Tsai

A new MOS current mode logic (MCML) style exhibiting capacitive coupling to enhance the switching speed of the digital circuits is proposed. The mechanism of capacitive coupling and its effect on the delay are analytically modeled. SPICE simulations to validate the accuracy of the analytical model have been carried out with TSMC 0.18 μm CMOS technology parameters. Several logic gates such as five-stage ring oscillator, NAND, XOR2, XOR3, multiplexer, and demultiplexer based on the proposed logic style are implemented and their performance is compared with the conventional logic gates. It is found that the logic gates based on the proposed MCML style lower the delay by 23 percent. An asynchronous FIFO based on the proposed MCML style has also been implemented as an application.

1. Introduction

The rapid advances in the VLSI technology have led to the development of high-resolution mixed-signal applications. These applications demand high performance digital circuits to be integrated with analog circuitry on the same chip. MOS current mode logic (MCML) style has been widely used in digital circuits design for mixed-signal applications as they provide an analog friendly environment due to the low switching noise [1–4]. MCML circuits exhibit high switching speed, high noise immunity and better power efficiency at high operating frequencies along with a drawback of static power consumption [5–8].

In mixed-signal applications, the digital circuits are extensively used in the realization of digital signal processor functional units such as finite impulse response (FIR) filter and FFT module. The functional units are required to perform computations at high speed to efficiently use the bandwidth in communication systems which is also increasing. Therefore it is necessary to improve the speed of conventional MCML circuits. In this paper, a new MCML style with capacitive coupling that increases the switching speed of the circuits is proposed.

The paper first presents a brief introduction to conventional MCML style in Section 2. Thereafter, the architecture of the MCML style with capacitive coupling is proposed in Section 3. The mechanism of capacitive coupling is explained, and an expression for the delay is derived. The theoretical propositions are validated through SPICE simulations using TSMC 0.18 μm CMOS technology parameters in Section 4. The simulation results of several logic gates and an asynchronous FIFO are also presented in the same section. Finally, the conclusions are drawn in Section 5.

2. Conventional MCML Circuits

A conventional MCML circuit consists of three main components which includes a pull-down network (PDN), a constant current source, and a load circuit. The circuit of a conventional MCML inverter is shown in Figure 1. Its PDN consists of a source-coupled transistor pair MC2-MC3 with input A. The constant current source MC1 generates the bias current I_{SS} while the load MC4 determines the output voltage swing [9]. The circuit works on the principle of current steering, where the bias current I_{SS} is steered to one of the circuit branches depending on the input voltage. When the voltage at input A is high, the bias current I_{SS} is steered to transistor MC2 and results in a low V_{OL} ($V_{OL} = V_{DD} - V_P$) at the output node Q and a high voltage ($V_{OH} = V_{DD}$) at the other node $\overline{Q}$ where V_P is the voltage drop across the load [9]. Conversely, when the input A is low, the bias current

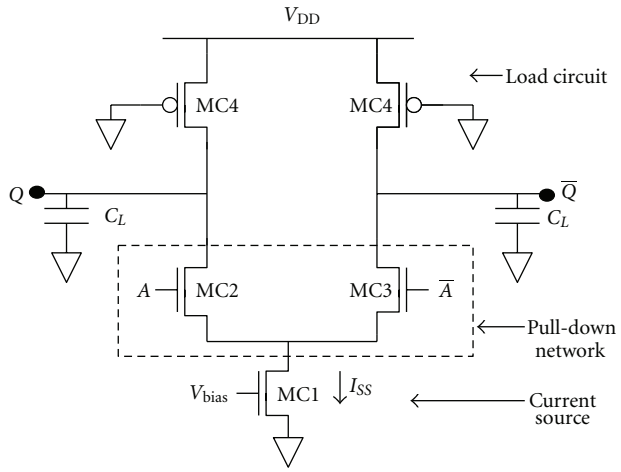


Figure 1: Conventional MCML inverter.

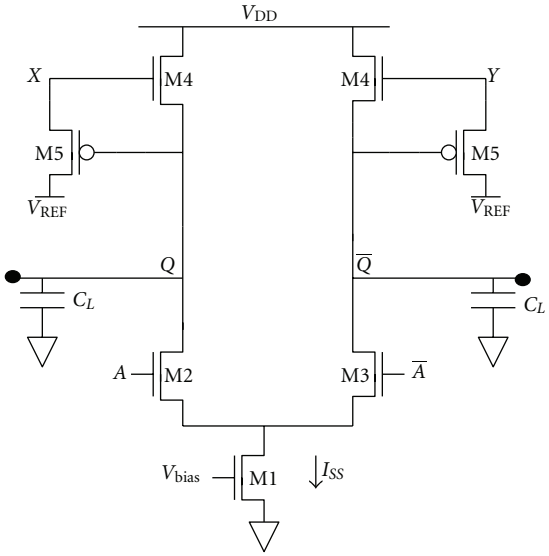


Figure 2: Proposed MCML inverter.

I_{SS} is steered to MC3, and a high voltage is obtained at the node Q.

3. Proposed MCML Style

3.1. Basic Architecture. The basic architecture of the proposed MCML style includes a PDN, a constant current source, and a load that exhibits capacitive coupling to increase the switching speed of the circuits. An inverter based on the proposed logic style is shown in Figure 2. Its PDN consists of a source-coupled transistor pair M2-M3 to implement the logic function and a current source M1 to generate the bias current I_{SS}. The load includes one NMOS and one PMOS transistors M4-M5 as shown in Figure 2. The reference voltage V_{REF} is chosen to be one threshold voltage above the supply voltage to allow the transistor M4 operation in linear region by making the potential at intermediate node

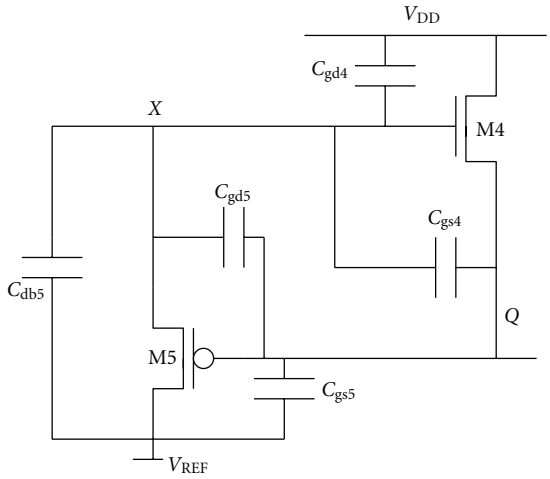


Figure 3: Parasitic capacitances of the transistors M4 and M5 in the load circuit.

(X, Y) as V_{REF}. The circuit also works on the principle of current steering. The phenomenon of capacitive coupling in the proposed circuit is described in the following section.

3.2. Capacitive Coupling. In this section, the phenomenon of capacitive coupling in the proposed circuit is described, and an expression for the delay is derived. An expression for conventional MCML inverter delay, using the same method, is also given.

3.2.1. Capacitive Coupling Analysis. The phenomenon of capacitive coupling occurs in the load during the transition at output node voltage. This can be explained by considering the half circuit of the inverter (Figure 2) and identifying the capacitances at different nodes in the load (M4-M5). The capacitances are shown in Figure 3 where C_{gdi}, C_{dbi} and C_{gsi} represent the gate-drain capacitance, drain-bulk capacitance and gate-source capacitance of the ith transistor in the load. The capacitances C_{gs4} and C_{gd5} can be represented by the coupling capacitance between node Q and the intermediate node X as C_C ($C_C = C_{gs4} + C_{gd5}$) whereas C_{gd4} and C_{db5} are the capacitance between node X and ground as C_X($C_X = C_{gd4} + C_{db5}$). Initially, let the voltage at input A is assumed to be equal to high so that the voltage at Q is low ($V_Q = V_{OL}$). For a high-to-low transition at the input, the transistor M2 turns off and the voltage at $Q(V_Q)$ begins to rise from V_{OL} to V_{OH}. This change in the output voltage gets coupled to node X through C_C.

Let i_{C_C} and i_{C_X} be the transient currents flowing through C_C and C_X, respectively. By applying the KCL, the current equation at node X can be written as

$$i_{C_C} - i_{C_X} \approx 0, \tag{1}$$

where it is assumed that negligible current flows in transistor M5. Substituting the current values

$$C_X \frac{dV_X}{dt} \approx C_C \frac{d(V_Q - V_X)}{dt} \tag{2}$$

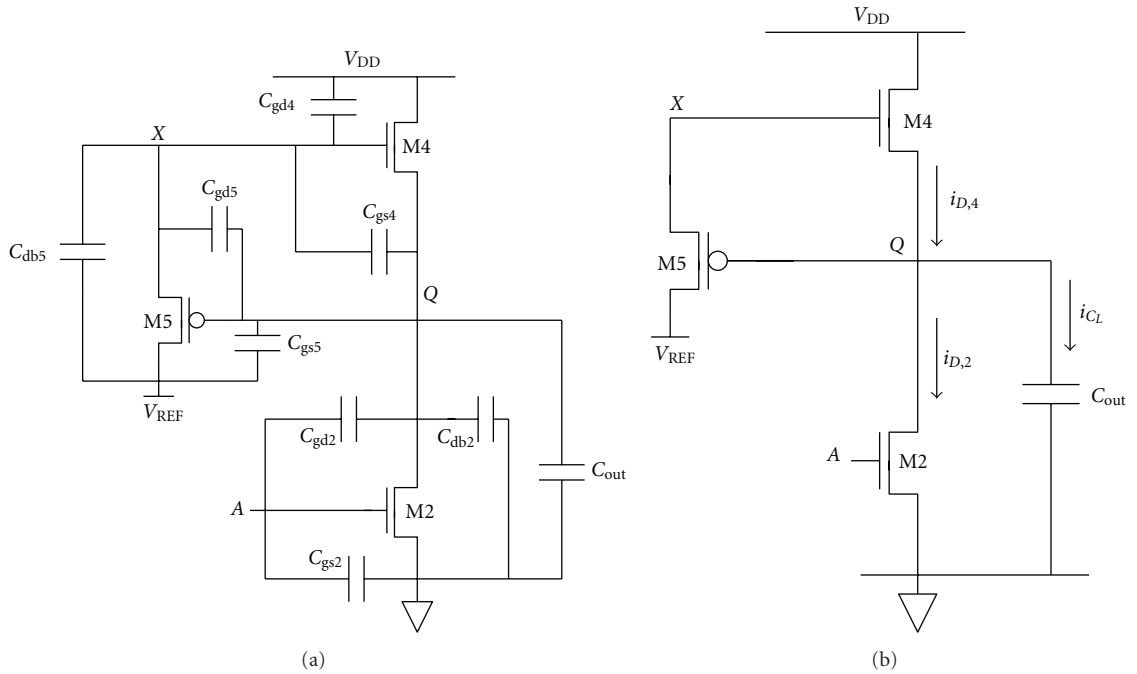


FIGURE 4: (a) Different capacitances in the half circuit of the proposed MCML inverter. (b) Half circuit of the proposed MCML inverter with load capacitance C_L.

which can be rearranged as

$$\frac{dV_X}{dt} = \frac{C_C}{C_C + C_X}\frac{dV_Q}{dt}. \tag{3}$$

It can be observe that an increase in the output voltage will result in an increase in the voltage of the node X. Multiplying (3) by dt and integrating both sides give

$$\int_{V_{REF}}^{V_X(t)} dV_X = \left(\frac{C_C}{C_C + C_X}\right)\int_0^t \frac{dV_Q}{dt}dt. \tag{4}$$

Completing the integral on the left side and rearranging it give

$$V_X(t) = V_{REF} + m\int_0^t \left(\frac{dV_Q}{dt}\right)dt, \tag{5}$$

$$= V_{REF} + m[V_Q(t) - V_Q(0)], \tag{6}$$

where $m = C_C/(C_C + C_X)$ represents the capacitance ratio and $V_{REF} = V_{DD} + V_T$. Further $V_X(t)$ may be expressed as

$$V_X(t) = \begin{cases} V_{DD} + V_T + m[V_Q(t) - V_{OL}] \\ \text{for } V_X(0) = V_{OL} \ldots \text{rising output} \\ V_{DD} + V_T - m\left[V_{OH} - V_Q(t)\right] \\ \text{for } V_X(0) = V_{OH} \ldots \text{falling output.} \end{cases} \tag{7}$$

3.2.2. Delay. The delay of the proposed MCML inverter can be modeled by solving the state equation of the output node in the time domain [10]. The half circuit of the proposed MCML inverter with the parasitic capacitances of the transistors is shown in Figure 4(a), where C_{out} represents the output that includes the interconnect capacitance and the input capacitance of the subsequent stage, respectively. The total capacitance at the output node C_L is shown in Figure 4(b) which may be computed as

$$C_L = C_{gd2} + C_{db2} + C_{gs5} + C_{gd5} + C_{gs4} + C_{out}. \tag{8}$$

The current through capacitor C_L may be written as

$$C_L\frac{dV_Q}{dt} = i_{C_L} = i_{D,4} - i_{D,2}. \tag{9}$$

When the input switches from high to low, the current through transistor M2 becomes zero such that C_L begins to charge through the load transistor M4. Thus, (9) reduces to

$$C_L\frac{dV_Q}{dt} = i_{D,4}. \tag{10}$$

The rising output voltage initiates the process of capacitive coupling in the load circuit. It is clear from (7) that the gate potential of M4 remains greater than or equal to V_{REF} during the charging process, hence M4 operates in linear region throughout the switching process. Therefore, the delay t_{PLH} may be calculated by solving (10) as

$$t_{PLH} = \int_{t_0}^{t_1} dt = C_L\int_{V_{OL}}^{V_{50\%}} \left(\frac{1}{i_{D,4}}\right)dV_Q. \tag{11}$$

Substituting $i_{D,4}$ in (11) results in

$$t_{\mathrm{PLH}} = \frac{2\,C_L}{k_{n,4}} \int_{\mathrm{V_{OL}}}^{\mathrm{V_{50\%}}} \left(\frac{1}{\left[2(\mathrm{V_X} - \mathrm{V_Q} - \mathrm{V_T})(\mathrm{V_{DD}} - \mathrm{V_Q}) - (\mathrm{V_{DD}} - \mathrm{V_Q})^2\right]} \right) \mathrm{dV_Q}, \tag{12}$$

where $k_{n,4}$ is the transconductance parameter of transistor M4

$$t_{\mathrm{PLH}} = \frac{2C_L}{k_{n,4}} \int_{V_{\mathrm{OL}}}^{V_{50\%}} \left(\frac{1}{(V_{\mathrm{DD}} - V_Q)[V_{\mathrm{DD}} + (2\,m - 1)V_Q - 2\,mV_{\mathrm{OL}}]} \right) dV_Q. \tag{13}$$

Evaluating the integral yields

$$t_{\mathrm{PLH}} = \frac{C_L}{k_{n,4}} \frac{1}{m(V_{\mathrm{DD}} - V_{\mathrm{OL}})} \times \ln \frac{V_{\mathrm{DD}} - V_Q}{V_{\mathrm{DD}} + (2\,m - 1)V_Q - 2\,mV_{\mathrm{OL}}} \Bigg|_{V_{\mathrm{OL}}}^{V_{50\%}}. \tag{14}$$

Equation (14) evaluates to

$$t_{\mathrm{PLH}} = \frac{C_L}{k_{n,4}} \frac{1}{m(V_{\mathrm{DD}} - V_{\mathrm{OL}})} \times \ln \frac{V_{\mathrm{DD}} + 2\,m(V_{50\%} - V_{\mathrm{OL}}) - V_{50\%}}{V_{\mathrm{DD}} - V_{50\%}}. \tag{15}$$

Similar analysis for a conventional MCML inverter (Figure 1) gives the delay as

$$t_{\mathrm{PLH}} = \frac{C_L}{k_{p,4}} \frac{1}{\left(V_{\mathrm{DD}} + V_{T,p}\right)} \left[\ln \frac{V_{\mathrm{DD}} + 2V_{T,p} + V_{50\%}}{V_{\mathrm{DD}} - V_{50\%}} - \ln \frac{V_{\mathrm{DD}} + 2V_{T,p} + V_{\mathrm{OL}}}{V_{\mathrm{DD}} - V_{\mathrm{OL}}} \right], \tag{16}$$

where $k_{p,4}$, and $V_{T,p}$ are the transconductance and the threshold voltage of PMOS transistor MC4, respectively.

4. Simulation Results

This section first demonstrates the phenomenon of capacitive coupling in the proposed MCML inverter (Figure 2) to verify the theoretical propositions presented in Section 3. Thereafter, the performance of the proposed MCML logic style is compared with the conventional MCML logic style by simulating different logic gates. Lastly, the simulation results for an asynchronous FIFO based on the proposed style are presented. All the simulations are performed by using TSMC 0.18 μm CMOS technology parameters with $V_T = 0.5\,V$ and a supply voltage of 1.8 V. The bias current of all the circuits is taken as 100 μA uniformly.

4.1. Proposed MCML Inverter. The proposed MCML inverter (Figure 2) with $m = 0.6$ and $V_{\mathrm{OL}} = 1.4$ V is designed and simulated to demonstrate the phenomenon of capacitive coupling. The aspect ratios of the transistors calculated are $(\mathrm{W/L})_{\mathrm{M1}} = 0.81/0.18$, $(\mathrm{W/L})_{\mathrm{M2,M3}} = 26.8/0.18$, $(\mathrm{W/L})_{\mathrm{M4}} = 0.68/0.18$, $(\mathrm{W/L})_{\mathrm{M5}} = 0.54/1.35$. The waveforms at the input, X, and output nodes are shown in Figure 5. It can be observed that the rising output voltage increases the voltage of node X which confirms the relation between voltages at node Q and X (3). Also, it can be observed that during a low-to-high transition at the output, for an output voltage $V_Q = 1.8$ V the potential at node X raises $V_X = 2.54$ V and is in accordance with (7) in Section 3.

The delay of the proposed MCML inverter expressed by (15) is validated for different values of m ranging from 0.1 to 1. The error in the predicted and the simulated delay is plotted for different values of m in Figure 6 and is always is less than 12%. Thus, the simulation results are in close agreement with the theoretical values.

4.2. Performance Comparison. Several logic gates such as inverter, five-stage ring oscillator (RO5), NAND, XOR2, XOR3, multiplexer, and demultiplexer based on the proposed MCML style and conventional MCML style are implemented for the purpose of comparison. The aspect ratios of the transistors in the PDN of the two styles are taken to be same. It may be noted that since all the circuits are implemented with the same supply voltage and bias current, therefore they will consume the same static power which is computed as the product of the supply voltage and bias current [4]. The delay of the gates based on both the logic styles is listed in Table 1. It can be observed that an improvement varying from 16% to 23% can be obtained in delay by using the proposed MCML logic style.

Table 1: Delay comparison of logic gates.

Logic gate	Delay of conventional MCML-based circuit (ps)	Delay of proposed MCML-based circuit (ps)	Percentage reduction
Inverter (C_L = 50 fF)	209	165	21%
Inverter (C_L = 100 fF)	419	340	18%
Inverter (C_L = 250 fF)	985	827	16%
Ring Oscillator (RO5)	290	250	14%
NAND	1795	1503	16%
XOR2	1950	1506	22%
XOR3	2153	1800	16%
2 : 1 Multiplexer	1955	1511	23%
1 : 2 Demultiplexer	1988	1600	19.5%

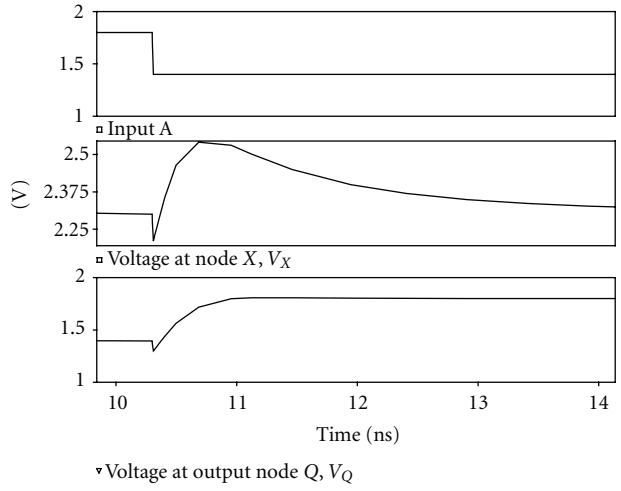


Figure 5: Voltages at different nodes of the proposed MCML inverter.

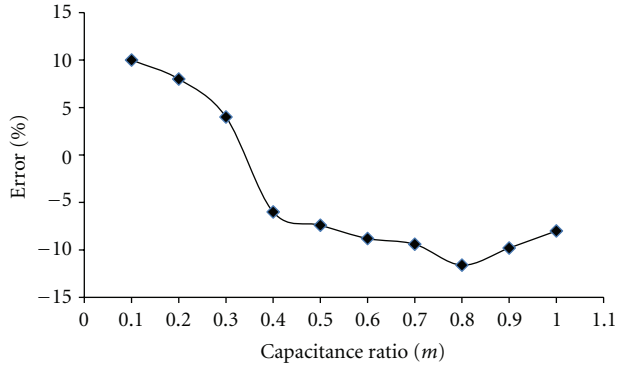


Figure 6: Error in the predicted and the simulated delay for the proposed MCML inverter with respect to m.

4.3. Application Example. An asynchronous FIFO is implemented as an application of the proposed MCML logic style. An asynchronous FIFO connects the sender and the receiver through a data bus consisting of separate request and acknowledge signals, and data signals [11–13]. The block diagram of a 4 stage FIFO is shown in Figure 7. It comprises of four stages wherein each stage consists of a functional unit and a control unit. The functional unit has a combinational stage for computing the result of each stage and a matched delay element inserted in the request line. The control unit employs a double-edge triggered flip-flop (DETFF) and a C element to control the communication between the successive stages. The signals shown as Req(in) and Ack(out) communicate the data, Data(in) between sender and the first stage. At the receiver side, the signals Req(out) and Ack(in) are used to synchronize the output data, Data(out) with the receiver and the last stage. Initially, the input data, Data(in) is loaded in the first stage of the FIFO and the Req(in) is asserted to low to start the data transfer. This results in a transition at the output of a C-element such that the data is stored in the DETFF of the first stage. At the same time an acknowledge signal Ack(out) is given to the sender. The stored data then flows through the different stages in the FIFO. Then, a request signal Req(out) is generated by the last stage to the receiver to enable the receiver to accept the data. This is followed by an acknowledge signal, Ack(in) from the destination to the last stage.

The circuits of the control unit elements based on the proposed logic style are shown in Figure 8. The waveforms obtained through the simulation of a four-stage asynchronous FIFO are shown below in Figure 9. The first three waveforms correspond to the input data Data(in), request signal (Req_in), and acknowledge signal (Ack_out) at the sender section. The last three graphs are the acknowledge signal Ack(in), data Data(out) and request signal Req(out). It can be found that the asynchronous FIFO based on the proposed logic style outputs the sampled data correctly.

5. Conclusion

In this paper, a new MCML style with capacitive coupling to increase the speed of the digital circuits is proposed. The phenomenon of capacitive coupling occurs in the load of the proposed style. The coupling phenomenon has been explained and analyzed. An expression for the delay has also been derived and validated through SPICE simulations using TSMC 0.18 μm CMOS technology parameters. Several logic gates based on the proposed logic style are implemented and their performance is compared with the conventional logic

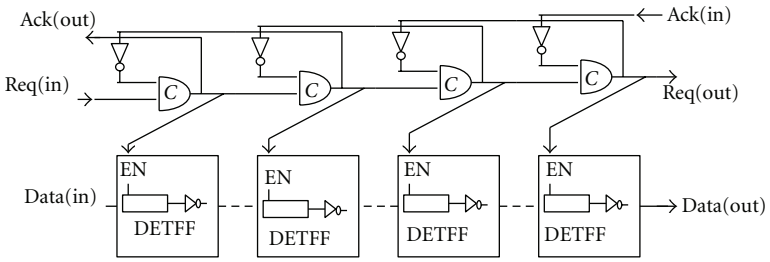


Figure 7: Block diagram of 4-stage asynchronous FIFO.

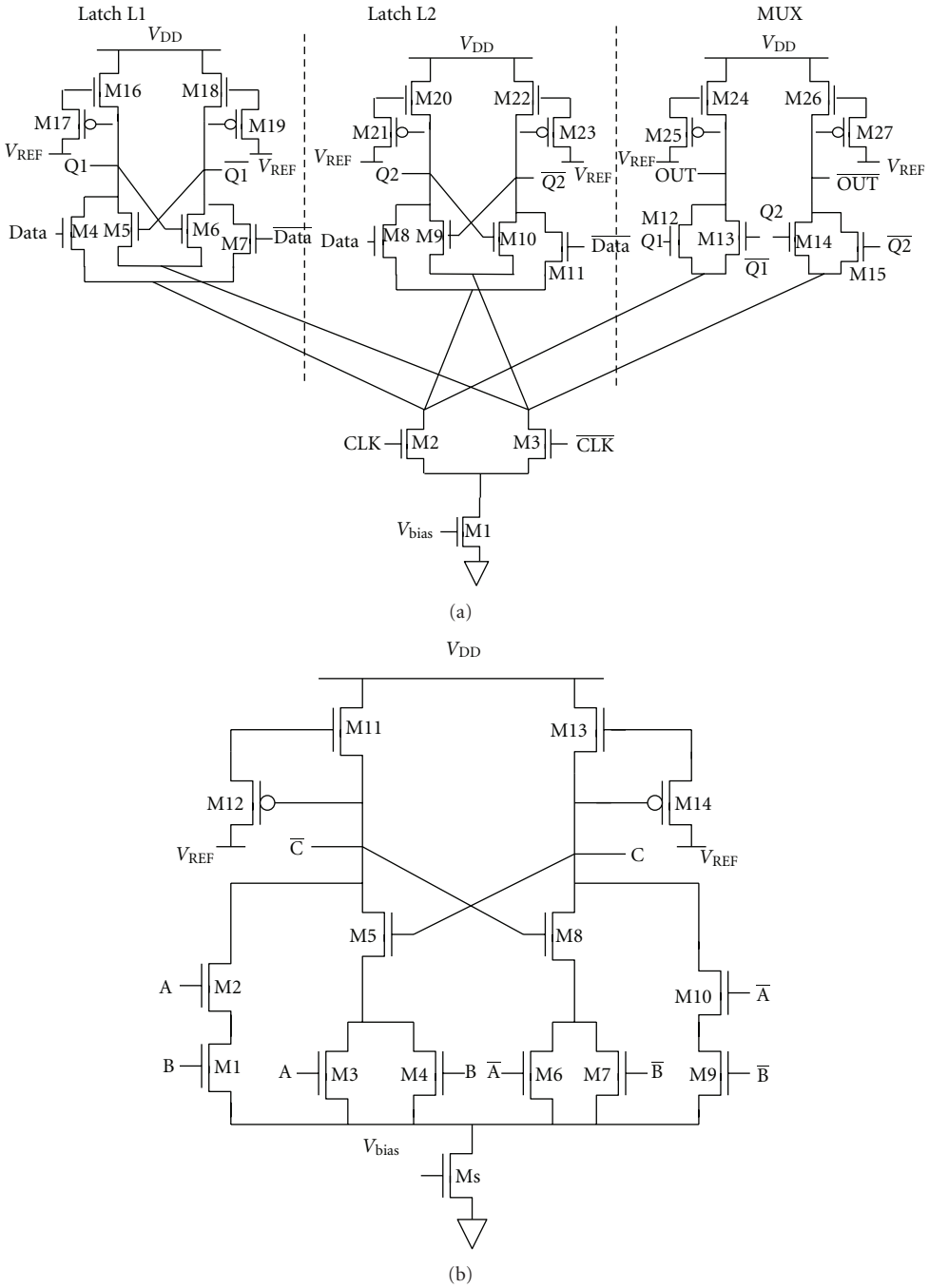


Figure 8: Control unit elements based on the proposed MCML style (a) DETFF (b) C-element.

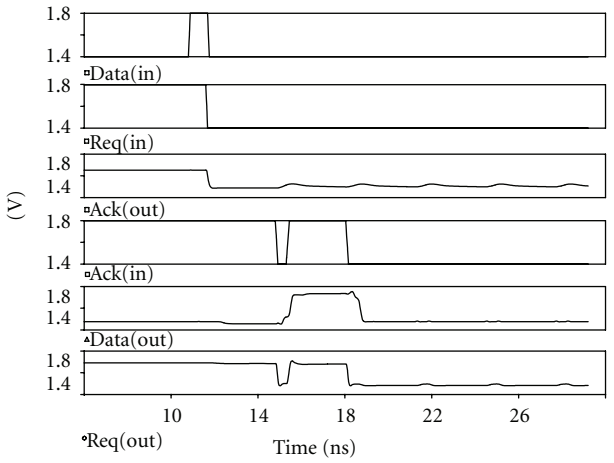


Figure 9: Transient response of the 4-stage asynchronous FIFO based on the proposed MCML style.

gates. It is found that a delay reduction of up to 23 percent delay reduction can be achieved by employing the proposed MCML style in the design of logic circuits.

References

[1] J. M. Musicer and J. Rabaey, "MOS Current Mode Logic for low power, low noise CORDIC computation in mixed-signal environments," in *Proceedings of the Symposium on Low Power Electronics and Design (ISLPED '00)*, pp. 102–107, July 2000.

[2] S. Bruma, "Impact of on-chip process variations on MCML performance," in *Proceedings of the IEEE Conference on Systems-on-Chip*, pp. 135–140, September 2003.

[3] M. Anis, M. Allam, and M. Elmasry, "Impact of technology scaling on CMOS logic styles," *IEEE Transactions on Circuits and Systems II*, vol. 49, no. 8, pp. 577–588, 2002.

[4] M. Alioto and G. Palumbo, *Model and Design of Bipolar and MOS Current-Mode logic (CML, ECL and SCL Digital Circuits)*, Springer, New York, NY, USA, 2005.

[5] O. Musa and M. Shams, "An efficient delay model for MOS current-mode logic automated design and optimization," *IEEE Transactions on Circuits and Systems I*, vol. 57, no. 8, pp. 2041–2052, 2010.

[6] H. Hassan, M. Anis, and M. Elmasry, "MOS current mode circuits: analysis, design, and variability," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 13, no. 8, pp. 885–898, 2005.

[7] J. B. Kim, "Low-power MCML circuit with sleep-transistor," in *Proceedings of the 8th IEEE International Conference on ASIC (ASICON '09)*, pp. 25–28, October 2009.

[8] T. K. Agarwal, A. Sawhney, A. K. Kureshi, and M. Hasan, "Performance comparison of static CMOS and MCML gates in sub-threshold region of operation for 32 nm CMOS technology," in *Proceedings of the International Conference on Computer and Communication Engineering*, pp. 284–287, May 2008.

[9] M. Alioto and G. Palumbo, "Power-delay optimization of D-latch/MUX source coupled logic gates," *International Journal of Circuit Theory and Applications*, vol. 33, no. 1, pp. 65–86, 2005.

[10] S. M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits: Analysis and Design*, Tata Mcgraw Hills, 3rd edition, 2006.

[11] M. Shams, J. C. Ebergen, and M. I. Elmasry, *Asynchronous Circuits*, John Wiley's Encyclopedia of Electrical Engineering, 1999.

[12] I. E. Sutherland, "Micropipelines," *Communications of the ACM*, vol. 32, no. 6, pp. 720–738, 1989.

[13] T. W. Kwan and M. Shams, "Design of multi-ghz asynchronous pipelined circuits in MOS current-mode logic," in *Proceedings of the 18th International Conference on VLSI Design: Power Aware Design of VLSI Systems*, pp. 301–306, January 2005.

IJSAA International Journal of Systems , Algorithms & Applications

Petri Net: History and its Connection to Fuzzy Sets and Linear Logic

Sangita Kansal
Department of Applied Mathematics,
Delhi Technological University,
Bawana Road, Delhi-110042 India
e-mail: sangita_kansal15@rediffmail.com

Abstract - Petri nets have been among the most succinct models that can describe the structure and dynamics of discrete event-driven systems. Various developments have taken place since the birth of Petri net. In this paper, efforts have been made to discuss these developments. The connection of Petri net to linear logic through fuzzy sets and lineale has also been discussed.

Keywords: *Petri net, Fuzzy set, Multi set, Lineale*

I. INTRODUCTION

Petri nets for modeling *concurrent* and *nondeterministic* computations have been an extensive area of research for more than four decades. The birth of Petri net was Petri's[6] dissertation in 1962, but most of the work is based on the final report of Holt [1] which translated Petri's dissertation into English as well as extended the work considerably.

Petri net basically consists of two sets namely, set of places and set of transitions. It also consists of markings which give the information about the number of tokens in a place. Petri's original nets did not allow multiple arcs between places and transitions. Further the firing rule was limited to requiring that a token reside in each input place to a transition and no token reside in the output place. A transition fired by removing the token from its inputs (which now become empty) and placing tokens in the outputs (which now become full). A transition could not fire if a token already resided in an output place. Thus a marking assigned either zero or one token to each place, $\mu: P \rightarrow \{0,1\}$.

Holt and Commoner [2] generalized the class of markings and the firing rule to allow arbitrary markings $\mu: P \rightarrow \{0,1,2,3..\}$.

One of the first formal definition was by *Patil [5]* where a Petri net was defined as a 4-tuple (P, T, F, μ^0) with P the set of places, T the set of transitions, $F \subseteq (P \times T) \cup (T \times P)$ called the flow relation and $\mu: P \rightarrow \{0,1,2,3..\}$ an initial marking. The arcs in the set F connected either a place with a transition or a transition with a place. There is no flow from transition to transition and place to place. The basic idea is that a flow between places/ transition and places/transitions cannot be direct and has to be mediated through a transition/ place.

Dangling arcs in the flow relation is not permitted in the sense that a place, if at all is connected, has to be connected to a transition.

But even in the concept of net discussed so far, the possibility that some places/transitions are isolated definitely exists, since the only structured constraint is that $F \subseteq (P \times T) \cup (T \times P)$. A tighter condition is thus frequently assumed, it is that $domF \cup codomainF = P \cup T$. This ensures that every place is connected to some transitions and every transition is connected to some place.

In other words, this means that no activity is possible without resources and no resource is such that it is neither consumed by an activity nor produced by an activity.

With all this in mind, *Petrinet was redefined as a 4-touple* (P, T, F, μ^0) *where,*

$P \cup T \neq \phi$ *and* $P \cap T = \phi$ *and*
$F \subseteq (P \times T) \cup (T \times P)$ *such that*
$domF \cup codomainF = P \cup T$.

This is the standard definition of petrinet.

After this, *Winskel [8]* generalized this definition by taking $F: (P \times T) \cup (T \times P) \rightarrow \mathbb{N}$ (where $\mathbb{N}$ is the set of natural numbers including 0). So Winskel allows multiple arcs between places and transitions. In this case $F(p, t)=n$ means that there are n arcs from a place p to transition t and means that there are m arcs from transition t to place p. This is by now a favorite version and is known as Winskel's definition.

Definition 1

A Petri net is a 4-touple $N = (P, T, F, \mu^0)$ *where*

i) P *is a nonempty set of places.*
ii) T *is the disjoint set of transitions.*
iii) F *is the multiset of* $(P \times T) \cup (T \times P)$*, called the casual dependency relation.*
iv) μ^0 *is a nonempty multiset set of places, called the initial marking which satisfies the restrictions.*
a) $\forall t \in T, \exists p \in P$ *such that* $F(p, t)>0$ *and* $\forall t \in T, \exists p \in P$ *such that* $F(t, p)>0$
b) $\forall p \in P, \mu^0(p) \neq 0$ *or* $(\exists t \in T$ *such that* $F(p, t) \neq 0)$ *or* $(\exists t \in T$ *such that* $F(t, p) \neq 0)$

Nets have a well-known graphical representation in which transitions are represented as boxes and places as circles with directed arcs between them. The initial marking is represented by placing 'tokens' to the appropriate multiplicity on each place.

Winskel [9] used Petri net in the context of category theory and several authors have investigated the connection between Petri net and linear logic.

Jensen[3] has also taken the flow relation as above and has given the definition in terms of the matrix form.

Definition 2

A Petri net is a 5-tuple $N = (P, T, I^-, I^+, \mu^0)$*, where*

i) P *is a nonempty set of places*
ii) T *is a nonempty set of transitions*

IJSAA International Journal of Systems , Algorithms & Applications

iii) $P \cap T \neq \phi$
iv) $\Gamma, I^+: P\times T \to \mathbb{N}$ *known as the negative and positive incidence functions respectively.*
v) $\forall p \in P$, $\exists t \in T$ *such that* $\Gamma(p, t) \neq 0$ *or* $I^+(p, t) \neq 0$*and* $\forall t \in T$, $\exists p \in P$ *such that* $\Gamma(p, t) \neq 0$ *or* $I^+(p,t) \neq 0$
vi) $\mu^0: P \to \mathbb{N}$ *is the initial marking.*

Here, $\Gamma(p, t)$ represents the number of arcs from p to t and $I^+(p, t)$ represents number of arcs from t to p. Γ, I^+ and μ^0 are viewed as matrices of size $|P|\times|T|$, $|P|\times|T|$, $|P|\times 1$ respectively.

Definition 3
A marking is a function $\mu: P \to \mathbb{N}$ *and a step is a function* $\sigma: T \to \mathbb{N}$*, i.e. markings are multiset of places and steps are multiset of transitions.*

Definition 4 *(Firing Rule)*
Let $N = (P, T, \Gamma, I^+, \mu^0)$ *be a Petri net. A transition* $t \in T$ *fires at a marking* μ *iff* $\Gamma(p, t) \leq \mu(p)$ $\forall p \in P$ *After firing at* μ*, t may occur, yielding a new marking* μ' *given by*

$$\mu'(p) = \mu(p) - \Gamma(p, t) + I^+(p, t) \quad \forall p \in P$$

In view of the above, the idea came into the mind of the author that, can the set of natural numbers $\mathbb{N}$ be replaced by the closed interval [0,1] (because there are number of connectives in [0,1] which can play the role of addition and subtraction)?

II. PETRI NET CORRESPONDING TO A FUZZY SET

If we change $\mathbb{N}$ by the closed unit interval [0,1], then we are in the domain of fuzzy sets. The definition is straight forward with $\mathbb{N}$ being replaced by [0,1]. However, when we come to the firing rule, the equation $\mu'(p) = \mu(p) - \Gamma(p, t) + I^+(p, t)$ which defines the action of t on the marking μ is open to a number of interpretations because of various counter parts available in [0,1] for addition and subtraction.

Consider [0, 1] with
$a \otimes b = min(1, a + b)$, $a^{\perp} = 1-a$,
$a \wp b = max(0, b - a)$
$a \ddot{-} b = \begin{cases} b - a & \text{if } b \geq a \\ 0 & \text{Otherwise} \end{cases}$
where $a, b \in [0, 1]$

By adapting the above structure on [0, 1], we can write the firing rule as

$$\mu'(p) = [\Gamma(p, t) \ddot{-} \mu(p)] \otimes I^+(p, t)$$
$$= [\Gamma(p, t)^{\perp} \wp \mu(p)] \otimes I^+(p, t)$$

then rule says: t fires at μ to remove $\Gamma(p, t)$ tokens if it can be done, then it adds $I^+(p, t)$ tokens subject to the condition that at no stage of the transition being in the firing situation, the number of tokens goes out of the interval [0,1].

If $\mu(p) \leq \Gamma(p, t)$ then this version does not say that the transition cannot fire, it says it removes all the tokens leaving the place empty and then adds $I^+(p, t)$ tokens in it. Thus a transition t can fire at any marking μ according to this version.

III. PETRI NET CORRESPONDING TO A LINEALE

Lineale was introduced by Hyland and Paiva in 1990. If we define a petrinet corresponding to a lineale by replacing $\mathbb{N}$ in the standard petrinet by a strong heyting lineale then this will be another link to linear logic as lineale is a model for linear logic.

For doing this it is necessary to give some definitions.

Definition 5
A pre-lineale is a structure $(L, \leq, o, e)$*, where*
i) $(L, \leq)$ *is a partially ordered set.*
ii) (L, o, e) *is a commutative monoid.*

The structures are compatible in the sense that, if $a \leq b$, we have $a \circ c \leq b \circ c$ $\forall c \in L$.

If there exists a largest $x \in L$ such that $a \circ x \leq b$ then this element is denoted $a \ddot{-} b$.

If in a pre lineale $a \ddot{-} b$ exists for all $a, b \in L$., then this structure is called lineale.

Definition 6
A strong heyting lineale is a structure $(L, \leq, o, e, \ddot{-}, —o, \perp)$ *where*
i) $(L, \leq, o, e, \ddot{-})$ *is a lineale.*
ii) $(L, —o, \perp)$*is a commutative monoid.*

Now we present the Petri net with respect to a strong heyting lineale L as follows:

A Petri net is a 5-tuple $N = (P, T, \Gamma, I^+, \mu^0)$*, where*
P and T are set of places and set of transitions respectively,
$\Gamma, I^+: P\times T \to L$, $\mu^0: P \to L$.
$\mu: P \to L$ *known as markings.*

By adapting the structure on a strong heyting lineale L, we can write the firing rule as

$$\mu'(p) = [\Gamma(p, t)^{\perp} —o \mu(p)] \circ I^+(p, t)$$

In this paper the basic idea was just to give the origin and history of Petri net. Though I have also given the connection of Petri net to linear logic through fuzzy sets and lineale to give an idea to the future researchers. Some of the work in this regard has already been published by Rai and Shukla[7] and Kansal and Kumar[4].

REFERENCES

[1] Holt, A ., Final report on the information system theory project, Technical ReportRADCTR-62-305, Rome air development centre, Griffiss air force base, New York (September 1968), 352 pages.
[2] Holt, A. and Commoner, F., Events and conditions , Applied Data Research,NewYork (1970).
[3] Jensoen, K., Coloured Petrinets, Lecture Notes in Computer Science, 254, (1986), pp 248-299.
[4] Kansal Sangita and Kumar, R., Fuzzysets, Linear logic and Petrinets, Journal of Interdisciplinary Mathematics, Vol 10(3), (2007), pp 369-381.
[5] Patil, S., Coordination of asynchronous events, Ph. D Dissertation, Deptt. Of Electrical Engg., M.I.T., (1970), 234 pages.
[6] Petri, C. A., Kommunikation mit automaten, Schriften des Institutes fur Instrumentelle Mathematik, Bonn, (1962).
[7] Rai Sangita, Shukla, W., and Srivastava, A. K., Fuzzy sets for linear logic, Comm. IFSA Math. Chapter, Vol 3 (1989), pp 55-58.
[8] Winskel, G., Categories of models for concurrency, Lecture notes in Computer Scinece, 197, (1984), pp 247-267.
[9] Winskel, G., Petrinets, algebras, morphisms and compositionality, Information and Computation, Vol 72(3), (1987), pp 197-238.

International Scholarly Research Network
ISRN Electronics
Volume 2012, Article ID 537469, 4 pages
doi:10.5402/2012/537469

Research Article

Regressive Structures for Computation of DST-II and Its Inverse

Priyanka Jain[1] and Anamika Jain[2]

[1] *Department of Electronics and Communication, Delhi Technological University, Shahbad Doulatpur, Main Bawana Road, Delhi-110042, India*
[2] *Department of Electronics and Communication, 245 Budhpur, GT Karnal Road, Industrial Area, National Highway 1, Delhi 110036, India*

Correspondence should be addressed to Priyanka Jain, priyajain2000@rediffmail.com

Received 31 August 2012; Accepted 14 October 2012

Academic Editors: P. Corsonello and J. Solsona

Efficient regressive structures for implementation of forward (DST-II) and inverse discrete sine transform (IDST-II) are developed. The proposed algorithm not only minimizes the arithmetic complexity compared to the existing algorithms (Wany (1990), Hupta and Rao (1990), Yip and Rao (1987), Murthy and Swamy (1992)) but also provides hardware savings over the algorithm (Jain et al. (2008)) by the same authors. The naturally ordered input sequence makes the new algorithms suitable for on-line computation.

1. Introduction

Discrete sine transform (DST) and discrete cosine transform (DCT) are used as key functions in many signal and image processing applications, for example, block filtering, transform domain adaptive filtering, digital signal interpolation, adaptive beam forming, image resizing, speech enhancement [1–6] and so forth, due to their near optimal transform coding performance. Both DST and DCT are good approximations to the statistically optimal Karhunen-Loeve transform [7, 8]. It is found that in case of image with high correlation coefficient, DCT-based coding results in better performance but for low correlation image, DST gives lower bit rate [8]. Since both DCT and DST are computationally intensive, many efficient algorithms have been proposed to improve the performance of their implementation [9–12]; however most of these are only good software solutions. Chiang and Liu [13] suggested a regressive structure for DCT-IV and DST-IV using second order digital filters. This paper is aimed at developing a similar regressive filter structure for DST-II/IDST-II that provides considerable hardware saving compared to the existing algorithm [14] and also reduces the computation complexity in terms of number of operation (multiplication and addition) as compared to the existing algorithms [15–18]. The advantage of this implementation is that no permutation is required for the input/output sequences, and therefore it is especially suitable for on-line computation.

2. Regressive Formulas of DST-II/IDST-II and Realization through Filter Structure

The DST of a sequence $y(n)$, $n = 1, 2, \ldots, N$ and its inverse are given by [14]

$$Y[k] = \gamma_k \sum_{n=1}^{N} y(n) \sin\left[\left(n - \frac{1}{2}\right)\frac{k\pi}{N}\right], \quad k = 1, 2, \ldots, N, \tag{1a}$$

where

$$\gamma_k = \sqrt{\frac{2}{N}}\varepsilon_k, \quad \varepsilon_k = \begin{cases} \dfrac{1}{\sqrt{2}} & k = N \\ 1 & k = 1, 2, \ldots, N-1, \end{cases} \tag{1b}$$

and

$$y(n) = \sqrt{\frac{2}{N}} \sum_{k=1}^{N} \varepsilon_k Y[k] \sin\left[\left(n - \frac{1}{2}\right)\frac{k\pi}{N}\right], \quad n = 1, 2, \ldots, N. \tag{2}$$

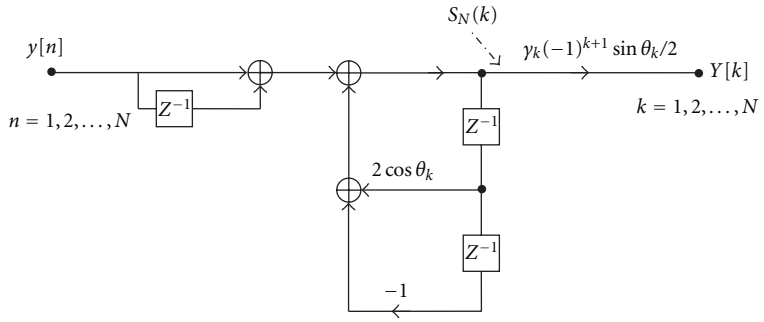


Figure 1: Regressive structure for computation of DST-II.

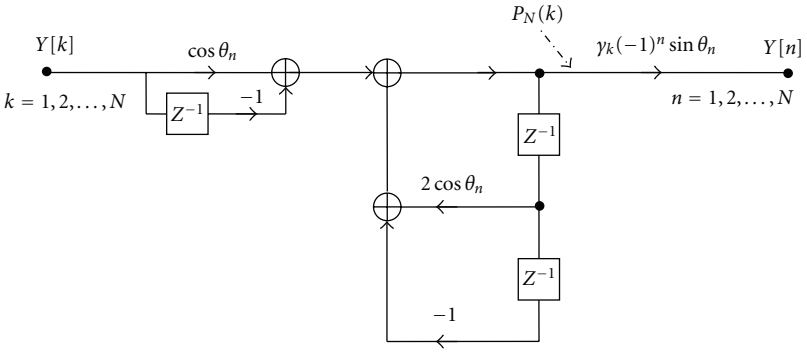


Figure 2: Regressive structure for computation of IDST-II.

Replacing n by $N - n$ in (1a), k by $N - k$ in (2), and after some algebraic manipulations, we obtain DST-II/IDST-II in the form

$$Y[k] = (-1)^{k+1}\gamma_k \sum_{n=0}^{N-1} y(N-n)\sin\left[\left(n+\frac{1}{2}\right)\frac{k\pi}{N}\right], \quad k = 1, 2, \ldots, N,$$

$$y(n) = (-1)^n\sqrt{\frac{2}{N}}\sum_{k=0}^{N-1}\varepsilon_k\, Y[N-k]\cos\left[\left(n-\frac{1}{2}\right)\frac{k\pi}{N}\right], \quad n = 1, 2, \ldots, N. \tag{3}$$

Let $\theta_k = k\pi/N$; $\theta_n = (n - 1/2)(\pi/N)$ and $x_n = \cos\theta_n$.

Further we define

$$V_n[x_k] = \frac{\sin(n+1/2)\theta_k}{\sin(\theta_k/2)}, \quad k = 1, 2, \ldots, N, \tag{4}$$

$$U_k[x_n] = \frac{\cos k\theta_n}{\cos\theta_n}, \quad n = 1, 2, \ldots, N. \tag{5}$$

We may rewrite (3) in the form

$$Y[k] = (-1)^{k+1}\sin\frac{\theta_k}{2}\gamma_k\sum_{n=0}^{N-1} y(N-n)V_k[x_k], \quad k = 1, 2, \ldots, N, \tag{6}$$

$$y[n] = (-1)^n\cos\theta_n\sqrt{\frac{2}{N}}\sum_{k=0}^{N-1}\varepsilon_k Y[N-k]U_k[x_n], \quad n = 1, 2, \ldots, N. \tag{7}$$

Further, we can write $V_{n+1}[x_k]$ as

$$V_{n+1}[x_k] = \frac{\sin(n+3/2)\theta_k}{\sin(\theta_k/2)}, \quad k = 1, 2, \ldots, N, \tag{8}$$

or

$$V_{n+1}[x_k] = \frac{2\sin(n+1/2)\theta_k\cos\theta_k - \sin(n-1/2)\theta_k}{\sin(\theta_k/2)}, \quad k = 1, 2, \ldots, N. \tag{9}$$

Equation (9) can be represented by a sinusoidal recursive formula as

$$V_{n+1}[x_k] = 2\cos\theta_k V_n - V_{n-1}, \quad k = 1, 2, \ldots, N. \tag{10}$$

In the same manner, we derive the sinusoidal formula for (5) as

$$U_{k+1}[x_n] = 2\cos\theta_k U_k - U_{k-1}, \quad n = 1, 2, \ldots, N. \tag{11}$$

In (6) and (7), we define

$$S_N(k) = \sum_{n=0}^{N-1} y(N-n)V_n[x_k], \tag{12}$$

$$P_N(k) = \sum_{k=0}^{N-1} Y[N-k]U_k[x_n]. \tag{13}$$

Using the sinusoidal recursive formula given in (10) and the fact that, $V_0 = 1$; $V_{-1} = -1$ and $U_0 = \cos\theta_n$, (12) can be written as

$$\begin{aligned} S_N(k) &= y(N)V_0 + \sum_{n=1}^{N-1} y(n-n)V_n[x_k] \\ &= y(N) + \sum_{n=0}^{N-2} y(N-n-1)V_{n+1}[x_k] \\ &= y(N) + \sum_{n=0}^{N-2} y(N-n-1)[2\cos\theta_k V_n - V_{n-1}] \\ &= y(N) + 2\cos\theta_k S_{N-1}(k) - y(N-1)V_{-1} - S_{N-2}(k). \end{aligned} \tag{14}$$

Similarly, (13) is written as

$$\begin{aligned} P_N(k) &= Y[N]U_0 + \sum_{k=1}^{N-1} Y[N-k]U_k[x_n] \\ &= \frac{Y[N]}{\cos\theta_n} + \sum_{k=0}^{N-2} Y[N-k-1]U_{k+1}[x_n] \\ &= \frac{Y[N]}{\cos\theta_n} + 2\cos\theta_k P_{N-1}(k) - Y[N-1] - P_{N-2}(k), \end{aligned} \tag{15}$$

where $U_{-1} = 1$. Using (6) and (12), the DST-II can be expressed as

$$Y[k] = (-1)^{k+1}\gamma_k \sin\frac{\theta_k}{2} S_N(k), \quad k = 1, 2, \ldots, N. \tag{16}$$

Similarly, using (7) and (13), the IDST expression is

$$y(n) = (-1)^n \gamma_k \cos\theta_n P_N(k), \quad n = 1, 2, \ldots, N. \tag{17}$$

From (16) and (17), DST-II/IDST-II can be implemented with the second order filter structure shown in Figures 1 and 2.

3. Performance

The proposed algorithm has been presented with the aim of realizing the DST/IDST of any length for on-line computation. It can be seen from the regressive structure for DST-II shown in Figure 1 that only one real multiplication per sample is required to bring the system in a state from which $S_N(k)$ can be computed. Similarly the regressive structure shown in Figure 2 for IDST shows that two real multiplications are required for $P_N(k)$ computation. The constant multiplication by $(-1)^{k+1}\gamma_k \sin(\theta_k/2)$ or $(-1)^n \gamma_k \cos\theta_n$ at the output need not be performed at every iteration of the difference equation but it is to be multiplied after the Nth step. Hence the total real multiplications for DST computation are $(N + 1)$ and for IDST computation are $(2N + 1)$. The number of additions for DTS as well as for IDST is $3N$. In Table 1 the number of multipliers, adders and latches required by the present algorithm is compared with those required in algorithms [14]. It can be seen from Table 1 that the new recursive algorithm is efficient in terms of saving the number of multipliers, adder, and latches compared to algorithm [14]. Both of the algorithms compute the DST/IDST of an N point real sequence. Although the proposed algorithm does not optimize the number of computational cycle, but has advantage in terms of hardware compared to previous approach [14]. Further, the filter structure is numerically stable, as it involves no division at all. Algorithm [14] needs addition and subtraction of input sequence whereas the input sequence for the present algorithm is in the natural order, which makes the proposed approach suitable for on-line computation. Table 2 shows the comparison of computational complexity of the proposed algorithm with the other algorithms [15–18]. It can be seen that the proposed algorithm not only reduces the number of multiplication significantly but also reduces the total number of operations.

Table 1: Hardware comparison for implementing DST-II/IDST-II coefficient.

S. No.	Comparison	Proposed		[14]	
		DST	IDST	DST	IDST
1	Latches	2	2	4	4
2	Real multiplier	2	3	4	4
3	Adder	3	3	6	8
4	Complex multiplier	0	0	0	0
5	Computation cycles	N	N	$N/2$	$N/2$

Table 2: Comparison of algorithms for computing DST in terms of addition and multiplication.

Algorithms	Number of addition $N = 2^M$	Number of multiplication $N = 2^M$
Fast DST, [15, 16]	$3MN/2 - N + 1$	$MN/2 + 1$
[17]	$4N + 2$	$6N$
[18]	$5N + 2$	$5N$
The proposed design	$3N$	$(N + 1)$

4. Conclusion

The proposed regressive algorithm has been derived mathematically and realized for kernels of DST and its inverse. This algorithm is effective for realization using software and hardware techniques. This algorithm provides substantial savings in actual number of multipliers, adders, and latches

in the hardware implementation required to perform the DST/IDST and reduces the arithmetic complexity compared to the existing algorithms. The proposed regressive algorithm has simple, regular, and modular filter structure and is particularly suitable for on-line computation.

Acknowledgment

The authors would like to thank the associate editor and referees for their valuable comments to improve the paper.

References

[1] Z. Wang, G. A. Jullien, and W. C. Miller, "Interpolation using the discrete sine transform with increased accuracy," *Electronics Letters*, vol. 29, no. 22, pp. 1918–1920, 1993.

[2] S. A. Martucci and R. M. Mersereau, "New approaches to block filtering of images using symmetric convolution and the DST or DCT," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS '93)*, pp. 259–262, May 1993.

[3] F. Beaufays, "Transform-domain adaptive filters: an analytical approach," *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 422–431, 1995.

[4] Y. S. Park and H. W. Park, "Arbitrary-ratio image resizing using fast DCT of composite length for DCT-based transcoder," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 494–500, 2006.

[5] L. Xueyao, X. Hua, and C. Bailing, "Noisy speech enhancement based on discrete sine transform," in *Proceedings of the 1st International Multi-Symposiums on Computer and Computational Sciences (IMSCCS '06)*, pp. 199–202, April 2006.

[6] S. C. Lim, D. Y. Kim, and Y. L. Lee, "Alternative transform based on the correlation of the residual signal," in *Proceedings of the 1st International Congress on Image and Signal Processing (CISP '08)*, pp. 389–394, May 2008.

[7] R. J. Clarke, "Relation between the Karhunen Loeve and Sine transform," *Electronics Letters*, vol. 20, no. 1, pp. 12–13, 1984.

[8] A. K. Jain, "Fast Karhunen-Loeve transform for a class of stochastic processes," *IEEE Transactions on Communications*, vol. 24, no. 9, pp. 1023–1029, 1976.

[9] V. Kober, "Fast algorithms for the computation of sliding discrete sinusoidal transforms," *IEEE Transactions on Signal Processing*, vol. 52, no. 6, pp. 1704–1710, 2004.

[10] D. Püschel and J. M. F. Moura, "Algebraic signal processing theory: cooley-Tukey type algorithms for DCTs and DSTs," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1502–1521, 2008.

[11] P. K. Meher and M. N. S. Swamy, "New systolic algorithm and array architecture for prime-length discrete sine transform," *IEEE Transactions on Circuits and Systems II*, vol. 54, no. 3, pp. 262–266, 2007.

[12] P. K. Meher, A. P. Vinod, J. C. Patra, and M. N. S. Swamy, "Reduced-complexity concurrent systolic implementation of the discrete sine transform," in *Proceedings of IEEE Asia Pacific Conference on Circuits and Systems (APCCAS '06)*, pp. 1535–1538, December 2006.

[13] H. C. Chiang and J. C. Liu, "A regressive structure for on-line computation of arbitrary length DCT-IV and DST-IV transforms," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 6, pp. 692–695, 1996.

[14] P. Jain, B. Kumar, and S. B. Jain, "Discrete sine transform and its inverse—realization through recursive algorithms," *International Journal of Circuit Theory and Applications*, vol. 36, no. 4, pp. 441–449, 2008.

[15] Z. Wang, "Fast discrete sine transform algorithms," *Signal Processing*, vol. 19, no. 2, pp. 91–102, 1990.

[16] A. Gupta and K. R. Rao, "Fast recursive algorithm for the discrete sine transform," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 3, pp. 553–557, 1990.

[17] P. Yip and K. R. Rao, "On the shift properties of DCT's and DST's," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 3, pp. 404–406, 1987.

[18] N. R. Murthy and M. N. S. Swamy, "On the computation of running discrete cosine and sine transform," *IEEE Transactions on Signal Processing*, vol. 40, no. 6, pp. 1430–1437, 1992.

Software Maintainability Prediction using Machine Learning Algorithms

Ruchika Malhotra[1] and Anuradha Chug[2]

[1]*Department of Software Engineering, Delhi Technological University, Delhi 110042, India*

[2]*University School of Information and Communication Technology, GGS IP University, Dwarka, New Delhi 110077*

ruchikamalhotra2004@yahoo.com, a_chug@yahoo.co.in

Abstract - **Software maintainability is one of the most important aspects while evaluating quality of the software product. It is defined as the ease with which a software system or component can be modified to correct faults, improve performance or other attributes or adapt to a changed environment. Tracking the maintenance behaviour of the software product is very complex. This is precisely the reason that predicting the cost and risk associated with maintenance after delivery is extremely difficult which is widely acknowledged by the researchers and practitioners. In an attempt to address this issue quantitatively, the main purpose of this paper is to propose use of few machine learning algorithms with an objective to predict software maintainability and evaluate them. The proposed models are Group Method of Data Handling (GMDH), Genetic Algorithms (GA) and Probabilistic Neural Network (PNN) with Gaussian activation function. The prediction model is constructed using the above said machine learning techniques. In order to study and evaluate its performance, two commercial datasets UIMS (User Interface Management System) and QUES (Quality Evaluation System) are used. The code for these two systems was written in Classical Ada. The UIMS contains 39 classes and QUES datasets contains 71 classes. To measure the maintainability, number of "CHANGE" is observed over a period of three years. We can define CHANGE as the number of lines of code which were added, deleted or modified during a three year maintenance period. After conducting empirical study, performance of these three proposed machine learning algorithms was compared with prevailing models such as GRNN (General Regression Neural Network) Model, ANN (Artificial Neural Network) Model, Bayesian Model, RT (Regression Tree) Model, Backward Elimination Model, Stepwise Selection Model, MARS (Multiple Adaptive Regression Splines) Model, TreeNets Model, GN (Generalized Regression) Model, ANFIS (Adaptive Neuro Fuzzy inference System) Model, SVM (Support Vector Machine) Model and MLR (Multiple Linear Regressions) Model which were taken from the literature. Based on experiments conducted, it was found that GMDH can be applied as a sound alternative to the existing techniques used for software maintainability prediction since it assists in predicting the maintainability more accurately and precisely than prevailing models.**

Keywords: **GMDH (Group Method of Data Handling), Genetic Algorithms, Probabilistic Neural Network (PNN), Software Maintainability, Software Maintainability Prediction Metrics and Modeling.**

1. INTRODUCTION

Software maintainability means the ease with which a software system or component can be modified to correct faults, improve performance or other attributes or adapt to a changed environment [1]. The change in the software is required to meet the changing requirements of customers which may arise due to many reasons such as change in the technology, introduction of new hardware or enhancement of the features provided etc. Producing software which does not need to be changed is not only impractical but also very uneconomical. This process of changing the software which has been delivered is called software maintenance. The amount of resource, effort and time spent on software maintenance is much more than what is being spent on its development. Thus, producing software that is easy to maintain may potentially save large costs and efforts. One of the main approaches in controlling maintenance cost is to monitor software metrics during the development phase. It is a matter of interest for researches to measure various attributes of software design in terms of inheritance, coupling, cohesion etc and predict its maintenance behaviour on the basis of their values. The problem of predicting the maintainability of software is widely acknowledged in the industry and much has been written on how maintainability can be predicted by using various tools and processes at the time of designing with the help of software design metrics [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16]. Studies have been conducted and found the strong link between Object Oriented software metrics and its maintainability. They have also found that these metrics can be used as predictors of maintenance effort. Accurate prediction of software maintainability can be useful because of the following reasons:

(a). It helps project managers in comparing the productivity and costs among different projects.

(b). It provides managers with information for more effectively planning the use of valuable resources.

(c). It helps managers in taking important decision regarding staff allocation.
(d). It guides about maintenance process efficiency.
(e). It helps in keeping future maintenance effort under control.
(f). The threshold values of various metrics which drastically affect maintainability of software can be checked and kept under control so as to achieve least maintenance cost.
(g). It enables the developers to identify the determinants of software quality so that they can improve design and coding.
(h). It helps practitioners to improve the quality of software systems and thus optimize maintenance costs.

To measure the maintainability we first find out the "change effort". It is defined as "how much amount of average efforts are required to add, change or delete existing classes". Software maintenance is very important as it consumes 70% of the time of any product's life and indeed it is very challenging. Despite this fact it is poorly managed because we really do not have good measures of software maintainability. The fundamental reality is stated by Morco [17] in the year 1982 in his book "Controlling Software Projects" that "you cannot control what you cannot measure". It really highlights the importance of a good measurement of software maintainability so that we can control it.

To measure the various features of object oriented paradigm such as inheritance, cohesion, coupling, memory allocation etc different metrics are carefully selected. We have studied various metrics available in literature and selected only those software metrics that have a strong relationship with software maintainability and used them while constructing our model for prediction of object oriented software maintainability. These metrics are Weighted Methods per Class (WMC), Depth of Inheritance (DIT), Number of Children (NOC), Lack of Cohesion in Methods (LCOM), Response For a Class (RFC), Message Passing Coupling (MPC), Data Abstraction Coupling (DAC), Number of Local Methods (NOM), and (SIZE1) traditional line of code, (SIZE2) total number of attributes and methods of a class. We divided our data into three parts. 60% of the data is used for training i.e. machine learn from the data patterns using specified algorithm, 20% of the data is used for validation and 20% of the data is used for testing. From the literature it is verified that this is the commonly accepted proportion used by most researchers and practitioners [6, 8, 9, 11, 12, 13, 14, 16, 18, 19].

The relationships between static software metrics and its maintainability is very complex and non linear, hence conventional statistical techniques based models, which are purely based on quantity, would not help much to the problem. Instead, use of machine learning algorithms to establish the relationship between metrics and maintainability would be much better approach as these are based on quantity as well as quality. Three machine learning algorithms were proposed and evaluated in this study. First proposed method is very powerful architecture called Group Method of Data Handling (GMDH). The main idea behind GMDH is that it tries to build a function (called a polynomial model) which would behave in such a way that the predicted value of the output would be as close as possible to the actual value of the output. The GMDH network is implemented with polynomial terms in the links and a genetic component to decide how many layers are to be built. The result of training at the output layer can be represented as a polynomial function of all or some of the inputs. Next proposed model was Genetic Algorithms (GA) which was based on the principles of Darwin's evolution theory. Over many generations, the "fittest" individuals tend to dominate the population. In predictions based problems, GA try to discover an optimal solution by simulating the evolution theory. For Predicting the object oriented software maintainability, the genetic algorithms start their job by first selecting a set of software metrics, which is constituted with collection of genes (solutions). GA then uses natural selection and genetics as a basis to search for the optimal gene and a set of software metrics that give the best classification rate. Third proposed model in this study is Probabilistic Neural Network (PNN) which is based on neural network. Neural network technology mimics the human brain's own problem solving process. As the human beings use their knowledge from earlier experiences to solve new problems or face situations, the neural network also considers earlier solved examples to create a scheme of "neurons" which makes new choices, classifications and predictions.

In this study we have compared performances of above mentioned machine learning algorithms with other well known algorithms applied in the last decade for the purpose of prediction of software maintainability such as GRNN (General Regression Neural Network) model, ANN (Artificial Neural Network) Model, Bayesian Model, RT (Regression Tree) Model, Backward Elimination Model, Stepwise Selection Model, MARS (Multiple Adaptive Regression Splines) Model, Tree Nets Model, GN (Generalized Regression) Model, ANFIS (Adaptive Neuro Fuzzy inference System) Model, SVM (Support Vector Machine) model and MLR (Multiple Linear Regressions) Model etc in terms of MRE (Magnitude of Relative Error), MMRE (Mean Magnitude of Relative Error), Pred(0.25) and Pred(0.75).

The rest of the paper is organized as follows: Section 2 highlights the objectives of the study; Section 3 provides overview of the related research work conducted on prediction of software maintainability. Section 4 describes the machine learning algorithms proposed in this study i.e. GMDH, GA and PNN along with their advantages. Section 5 summarizes well thought-out selection of software design metrics and source of data considered in this empirical study. Section 6 includes the experimental setup, Results of the study, Analysis of results and Discussion. In Section 7, threats to validity have been discussed and finally Section 8 concludes the paper.

2. STUDY OBJECTIVE

The biggest irony of the software industry is that the largest cost associated with any software product over its lifetime is actually its maintenance cost. Most suggested approaches by all researchers for controlling maintenance costs is to utilize software metrics during the development phase. Studies examining the link between OO software metrics and maintainability have found that in general these metrics can be used as predictors of maintenance effort [6, 8, 9, 11, 12, 13, 14, 16]. The result shows in almost all the studies that the prediction accuracy of one model is more accurate on one dataset but is less accurate for another dataset. Although a number of maintainability prediction models have been developed in last two decades, they have low prediction accuracies according to the criteria suggested by Conte *et al.* [20]. Therefore, it is necessary to explore new techniques, which are not only easy in use but also provide high prediction accuracy for the purpose of maintainability prediction.

The GMDH algorithm [21, 22] is ideal for complex, unstructured system where the investigator is only interested in obtaining a high order input-output relationship [23]. Also, the GMDH algorithm is heuristic in nature and not based on solid foundation as is regression analysis. For many end users it may be more convenient to have such a model, which is able to make predictions using familiar polynomial formulas which are widely understood. GMDH is formulated as neural network architecture, and is called a polynomial network however; the output of the model is in the form of standard polynomial function. In fact, the GMDH network is not like regular feed forward networks and was not originally represented as a neural network. The GMDH algorithm and its modified versions have been previously applied to wide array of problems to ascertain predictions [23]. In the year 2009, it was also used for the prediction for software reliability [24] where it was proved as one of the best available models. In the current study an attempt was made to apply this model perhaps for the first time for the task of software maintainability prediction using Object Oriented software design metrics. The background of proposing GMDH model in software maintainability was that if it had proven empirically to predict reliability with least errors compared with other techniques [24], than possibly it may be useful as a sound alternative to existing techniques for maintainability predictions. Moreover, it was presumed that better cohesion and benefit to industry would be gained, if the same model can effectively predict both reliability and maintainability of newly developed software since they work in tandem to achieve the overall goal of software quality. The objective of our study was to apply GMDH model along with two other models GA and PNN; all three are machine learning algorithms and compare them with prevailing prediction models proposed in last decade to ascertain their performance in software maintainability.

3. RELATED WORK

There are several models and metrics proposed in literature to predict the maintainability of the softwares. These methods vary from simple statistical models such as regression analysis to complex machine learning algorithm such as neural networks etc. Various methods proposed in the literature for the prediction of maintainability have been summarized in Table 1. We elaborate few important studies here. Multiple Linear Regression (MLR) Model was used by Li and Henry to predict maintenance effort in 1998 [25] in which they not only created MLR model for prediction but also successfully earmarked those metrics which have strong impact on prediction of object oriented software maintainability. In the year 2000 Muthanna *et al.* also used polynomial regression to establish the relationship between design level metrics [26] and the corresponding maintainability of Industrial software. The results have shown that predicted values using polynomial regression were quite close to actual values. Dagpinar *et al.* also based their study on empirical data to establish the relationship between software metrics and its maintainability however instead of design level metrics of structure languages, the metrics were replaced by object oriented metrics. They recorded significant impact of two metrics i.e. direct coupling metric and size metric on software maintainability while other parameters like cohesion, inheritance and indirect coupling were not considered significant by them [18]. Fioravanti and Nesi in 2001 [5] presented a metric analysis to identify which metrics would be better ranked for its impact on prediction of adaptive maintenance for object-oriented systems.

The model and metrics proposed have been validated against real data by using MLR (Multilinear Regression Analysis) Model. The validation has identified that several metrics can be profitably employed for the prediction of software maintainability. Misra used linear regression in 2005 [7] and presented an empirical study which was based on intuitive and experimental analyses. It used a suite of twenty design and code measures to obtain their indications on software maintainability. Thwin and Quah [8] used neural networks to build object oriented software maintainability prediction models. Koten and Gray [11] used Bayesian Belief Network (BBN) to predict object-oriented software maintainability in year 2006. Many researchers [6, 8, 9, 11, 12, 13, 14, 16] have used two datasets (UIMS and QUES) proposed by Li and Henry [2] for model building and its evaluation. Zhou and Leung [12] used Multivariate Adaptive Regression Splines (MARS) for predicting object-oriented software maintainability in year 2007. They compared the performance of the MARS model with other very popular four models Multivariate linear regression (MLR), Support Vector Regression (SVR), Artificial Neural Network (ANN) and Regression Tree (RT). The results provided by them [12] suggest that MARS can predict maintainability more accurately and precisely with less error when compared with other models for the QUES dataset, and as accurate as the best model for the UIMS dataset. In the last decade some machine learning algorithms have also been proposed and evaluated. It has been verified empirically that the machine learning algorithms can predict maintainability more accurately and precisely. Aggarwal *et al.* suggested the use of Fuzzy model; Kaur *et al.* [14] stated the use of soft computing approaches such as ANN, FIS and ANFIS. Elish *et al.* [13] used Tree Nets and proved that they also provide competitive results when compared with other models. Recently Cong and Liu [19] have used Support Vector Machine. They conducted their study where the code was written to implement detection for the "temper proof HTML web page" that can be either used standalone or embedded as component within other layers of applications. The code was written in C++. Hidden Markov Model (HMM) was used by Ping [15] to define health index of a product in literature and suggested that it works as a weight on the process of maintenance behaviour over a period of time. The detail classification is given in Table 1.

TABLE 1
METHODS CLASSIFICATION

S.No.	Author	Year	Methods Used
1.	G. M. Berns.	1984	Maintainability Analysis Tool (Kind of Lexical Analyzer)
2.	D Kafura and R Reddy	1987	Static Analyzers (To count)
3.	Wake, S. and S. Henry	1988	Multiple Linear Regression Model
4.	Li W., Henry S.	1993	A Classic Ada Metric Analyzer [Based on LEX and YACC of UNIX
5.	RD Banker, S M Datar, CF Kemerer and Dani Zweig	1993	Statistical Model which assign weight to each metric
6.	F Niessink,HV Vliet	1997	Regression models
7.	D Stavironoudis, M Xenos, D Christodolakis,	1999	Experts Judgements
8.	F Fioravanti, PNesi	2001	Multiple linear Regression Model
9.	K.K. Aggarwal, Y Singh, JK Chhabra	2002	Fuzzy Model
10.	M Dagpinar, JH. Jhanke	2003	Multivariate Analysis, Correlation, Best subset Regression Model
11.	LA Stamelos, DE Sakellaris	2003	Bayesian belief networks

12.	SC Misra	2005	Linear Regression, co relation and multiple regression
13.	MMT Thwin , Quah,	2005	General Regression Neural Network (GRNN)
14.	K.K. Aggarwal, YSingh, P Chandra and M Puri	2005	Fuzzy Model
15.	K.K. Aggarwal, YSingh, A Kaur, and R Malhotra	2006	Artificial Neural Network
16.	C.V Koten, A.R. Gray	2006	Of Bayesian network, Neural network
17.	Y Zhou, H Leung	2007	Multiple adaptive regression splines (MARS)
18.	N.N Prasanth, S.Ganesh, G. Dalton	2008	Fuzzy repertory table (FRT) and Regression analysis
19.	N.N Prasanth, S.P.Raja, X.Birla, K.Navaz, SAA Rahuman	2009	Used Static Analysers and set threshold values of these metrics
20.	WANG Li-jin, HU Xin-xin, NING Zheng-yuan KE Wen-hua	2009	Projection Pursuit Regression (nonparametric multivariate regression technique)
21.	MO. Elish and KO Elish	2009	TreeNets
22.	A Kaur, K Kaur, R Malhotra	2010	Artificial Neural Network, Fuzzy Inference System (FIS), Adaptive Neuro Fuzzy Inference System (ANFIS)
23.	L Ping	2010	Hidden Markov Model (HMM) is used to simulate the maintenance behaviors shown
24.	C Jin, JA Liu	2010	Support Vector Machine
25.	F Marzoughi, MM Farhangian, A Marzoughi, ATH Sim	2010	Bayesian network theory

4. PROPOSED MODELS

This section is further divided into three sub-sections in which all the three proposed models are discussed in detail:

1. Group Method of Data Handling (GMDH)
2. Genetic Algorithm (GA)
3. Probabilistic Neural Network (PNN) with Gaussian Activation Function

4.1 Group Method of Data Handling (GMDH)

Russian Scientist A.G. Ivakhnenko introduced a technique in 1966 [21, 22], for constructing an extremely high order regression type model termed as GMDH. The algorithm, GMDH builds a multinomial of degree in hundreds, whereas standard multiple regression Boggs down in computation and linear dependence. The GMDH model has been described in Section 2 however few inherent advantages with GMDH approach are briefly highlighted as under:

(i). GMDH can predict the outcome even with smaller training sets.

(ii). The computational burden is reduced with GMDH model.

(iii). The procedure automatically filters out input properties that provide little information about location and shape of hyper surface.

(iv). A multilayer structure maintained in GMDH model is a computationally feasible way to implement multinomial of high degree.

The GMDH model has a forward multi-layer neural network structure. Each layer consists of one or more units

wherein two inputs arcs and one output arc is attached with every unit. Each unit corresponds to Ivakhnenko polynomial form [21, 22].

$$Z = a + bx + cy + dx^2 + exy + fy^2 \quad (1)$$

Or

$$Z = a + bx + cy + dxy \quad (2)$$

Where variables x and y are input variables and Z is output variable and a, b, c ...f are the parameters. Fujimoto et al. [27] described that the basic technique of GMDH learning algorithm is a self-organization method and it fundamentally consists of the following steps:

(i). Given a learning data sample including a dependent variable y and independent variables x1, x2, ... , xm ; split the sample into a training set and a checking set.

(ii). Feed the input data of m input variables and generate combination (m, 2) units from every two variable pairs at the first layer.

(iii). Estimate the weights of all parameters 'a' to 'f' in formula as per either using equation (a) or equation (b). Applying it on training dataset in the next step. In this study, stepwise regression method with formula given in equation (b) was employed.

(iv). Compute mean square error between actual and predicted value of each unit.

(v). Sort out the unit by mean square error in decreasing order and eliminate bad units.

(vi). Set the prediction of units in the first Layer to new input variables for the next layer, and build up a multi-layer structure by applying Steps (ii) (v).

(vii). When the mean square error become larger than that of the previous layer, stop adding layers and choose the minimum mean square error unit in the highest layer as the final model output. Steps (iv) and (v) describe an important and basic technique of GMDH algorithm. It is called regularity criteria and leads to achieving the minimum error at Step (vii) [22].

Some GMDH networks can be large. When it has around 7 or more input variables, we have to separate it into several sub-networks with 6 or less input variables since the rule extraction process becomes too complex with many input variables. Using GMDH network structure developed from the dataset of nominal dependent variable y and independent variables x the following model is obtained [27]:

Output Layer:

$$X_1(k) = a_1(k) + b_1(k) \cdot x_1(0) + c_1(k) \cdot x_2(0) + d_1(k) \cdot x_1(0) \cdot x_2(0) \quad (3)$$

Hidden Layers

$$X_2(k) = a_2(k) + b_2(k) \cdot x_1(0) + c_2(k) \cdot x_3(0) + d_2(k) \cdot x_1(0) \cdot x_3(0) \quad (4)$$

Input Layer

$$X_3(k) = a_3(k) + b_3(k) \cdot x_2(0) + c_3(k) \cdot x_4(0) + d_3(k) \cdot x_2(0) \cdot x_4(0) \quad (5)$$

Where G is the maximum layer, $X_i(k)$ is the output, $a_i(k)$ ..... $d_i(k)$ are the parameters of unit i in layer k and $x_n(0)$ is input variable n. After rule extraction for each sub-network of GMDH, several sub-rules are obtained. The action part of the sub-rule for the lower sub-network is corresponds to the condition part of the sub-rule for the upper sub-network. Then, by integrating all of the sub-rules, we can obtain a rule set for the whole GMDH network as shown in Figure 1.

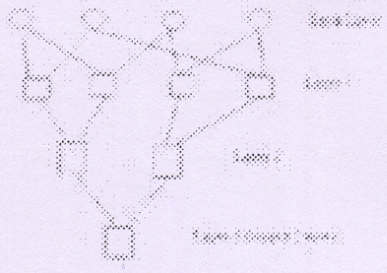


FIGURE 1 : STRUCTURE OF MULTI-LAYER NETWORK

4.2 Genetic Algorithms (GA)

A Genetic Algorithm is an adaptive system motivated by biological system proposed in Charles Darwin's evolution theory. It is a high level simulation. The GA starts with a set of solutions (represented by chromosomes) called population. GA is a search heuristic and it mimics the process of natural evolution. This heuristic is routinely used to generate useful solutions for optimization and search problems. Best solutions from one population are then taken and used to form a new population which will be better than the old one. While choosing the solutions, their fitness function is evaluated. Those solutions which are more close to fitness function have more probability to be selected. We say that the more suitable solutions have more chances "to survive". This process is repeated until some condition is satisfied such as achievement of best solution. Hence the population is improved over generations to accomplish the best solution. Indeed, GA is the methods designed to optimize the solutions

of prediction problem by simulating the "evolution behavior". The following processes are repeatedly applied until an optimized solution to the given problem is found:

(i). Natural selection
(ii). Crossover
(iii). Mutation

When this process is repeated over time, the better-fit individuals are the ones who survived; hence the genetic algorithms are also called as function optimizers. While implementing GA we first create a population with or without fixed size; First time usually, this population is randomly generated. Each individual of this population is then tested against "fit function". Reproductive opportunities are given to those individuals who have a better solution to the target problem and they have better chances of survival. Those individual solutions of the populations which are poorer and produce "weaker" solutions, they have less chances of survival. The "goodness" of a solution is defined in terms of the problem which needs to be solved. While solving any issue using GA the researchers first break the given issue into two problems i.e. the encoding problem and the evaluation problem. While designing the evaluation function, utmost care has to be kept in mind. The entire algorithm would fail if the evaluation function is poorly designed. It should be capable of measuring correctly the solution to the given problem. Evaluation function need not to be a mathematical expression and it could be a complete simulation.

The following are general steps implemented when using GA algorithms:

(i). Generate a random initial population.
(ii). Create the new population by applying the selection and reproduction operators to select pairs of strings. The number of pairs will be the population size divided by two, so the population size will remain constant between generations.
(iii). Apply the crossover operator to the pairs of the strings of the new population.
(iv). Apply the mutation operator to each string in the new population.
(v). Replace the old population with the newly created population.
(vi). Copy the best-fitted individual(s) to the newly created population to warrantee evolution.
(vii). (If the number of iterations is less than the maximum go to step two, else stop) OR (If the fitness of the best result does not get better over certain number of iteration, then stop).

4.3 *Probabilistic Neural Networks (PNN)*

This network has been originated from Neural Networks [16, 28, 37]. In the neural networks, it looks for patterns in training sets of data, learn these patterns, and develop the ability to correctly classify new patterns or to make forecasts and predictions. Neural networks excel at problem diagnosis, decision making, prediction, and other classifying problems where pattern recognition is important and precise computational answers are not required. Neural network starts its job by first recognizing patterns and trains the network. Training continues until the network reaches the conditions set in the 'Training' and 'Stop Training Criteria' module. This module calls different learning sub programs depending upon the paradigm and architecture we select. PNN is a feed forward neural network created by Specht [16] around 1990. It is based on Bayesian network and Kernel Fisher discriminate analysis. In a PNN, the operations are organized into a multilayered feed forward network with four layers:

(i). Input layer
(ii). Hidden layer
(iii). Pattern layer/Summation layer
(iv). Output layer

First layer is input layer where one neuron is present for each independent variable. The next layer is the hidden layer. This layer contains one neuron for each set of training data. It not only stores the values of the each predictor variables but also stores each neuron along with its target value. Next is the Pattern layer. In PNN networks one pattern neuron is present for each category of the output variable. Last layer is output layer. At this layer weighted votes for each target category is compared and selected. PNN are known for their ability to train quickly on sparse datasets as it separates data into a specified number of output categories. The network produces activations in the output layer corresponding to the probability density function estimate for that category. The highest output represents the most probable category. In the proposed study the number of neurons in the input layer at Slab 1 is equal to the number of inputs in our problem i.e. we have selected 11 independent variables summarized in Table 2 as inputs. The number of neurons in the output layer i.e. Slab 4 corresponds to the number of outputs. In the proposed study "Change" is taken as output variable. The number of neurons in the hidden layer defaults to the number of patterns in the training set because the hidden layer consists of one neuron for each pattern in the training set. We inspected smoothing factor for each link and apply it to all links. The smoothing factor that is defined during the design stage is default but we changed it in the training

sessions in order to make predictions more accurate and precise. We have experimented with different smoothing factors to discover which works best for our problem. We have applied the trained network to our training set, and to a test set too, using different smoothing factors and selected which was giving us the best answers. Then we used this module to train PNN networks. Unlike back propagation networks, which require feedback of errors and subsequent adjustment of weights and many presentations of training patterns, training a PNN network is very fast because it requires that each pattern be presented to the network only once during training. During the training session we can see the number of learning events completed during training which is also called as "epoch". Training can be done in real time since training is almost instantaneous. When data is sparse, training is superior to other network types. The success of PNN networks is dependent upon the smoothing factor. The adaptive PNN network is very powerful as during the building of neural networks, it uses genetic algorithms. Initially, we developed a GA based network algorithm that uses the GA directly with Calibration to improve the network's generalization. There are three ways for calibration of PNN which are as under:-

(i). Iterative Calibration Proceeds in Two Parts. The first part trains the network with the data in the training set. The second part uses Calibration to test a whole range of smoothing factors, trying to hone in on one that works best for the network created in the first part.

(ii). Genetic Adaptive: Uses a genetic algorithm to find appropriate individual smoothing factors for each input as well as an overall smoothing factor. The input smoothing factor is an adjustment used to modify the overall smoothing factor to provide a new value for each input.

(iii). None: In this calibration technique simply trains the network and we do not find an overall smoothing factor. The value for the smoothing factor is default chosen and applied. The user will have to manually adjust the smoothing factor by entering a new one in the edit box while using this module.

Even though PNN are slower and require more memory space, there are several advantages of PNN such as they are much faster, more accurate, and relatively insensitive to outliers, use Bayes optimal classification approach and generate accurate predicted target probability.

5. METRICS AND DATA

First we define the goal of this empirical study which is as follows:

Model : Evaluate GMDH model, Genetic model and PNN model (with Gaussians activation function) for the purpose of predicting object oriented software maintainability with respect to its prediction accuracy against the prevailing models like GRNN model, ANN Model, Bayesian Model, MARS Model, TreeNets, SVM model, Generalized Regression Model and ANFIS Model proposed by various researchers and practitioners during previous decade.

Metrics: We have worked on the set of metrics initially proposed by Chidamber et al. [29], and later by Li and Henry [2, 25, 30] and revised in Aggarwal et al. [36] as given in Table -2.

TABLE 2
METRICS DEFINITION

Metrics	Definition
WMC (Weighted Methods per Class)	The sum of McCabe's cyclomatic complexities of all local methods in a class. Let a class K1 with method M1…… Mn that are defined in the class. Let C1…….Cn be the complexity of the methods. W $WMC = \sum_{i=1}^{n} Ci$
DIT (Depth of Inheritance Tree)	The depth of a class in the inheritance tree where the root class is zero.
NOC (Number of Children)	The number of child classes for a class. It counts number of immediate sub classes of a class in a hierarchy.
RFC (Response For a Class)	The number of local methods plus the number of non local methods called by local methods.

LCOM (Lack of Cohesion of Methods)	The number of disjoint sets of local methods. Each method in a disjoint set shares at least one instance variable with at least one member of the same set.
MPC (Message Passing Coupling)	The number of messages sent out from a class.
DAC (Data Abstraction Coupling)	The number of instances of another class declared within a class.
NOM (Number of Methods)	The number of methods in a class.
SIZE1 (Lines of code)	The number of lines of code excluding comments.
SIZE2 (Number of properties)	The total count of the number of data attributes and the number of local methods in a class.
CHANGE (Number of lines changed)	The number of lines added and deleted in a class, change of the content is counted as two.

Datasets: In our study we use two most popular object-oriented maintainability datasets which are also published by Li and Henry [2]: UIMS and QUES datasets. Their Descriptive statistics is given in Table -3 and Table - 4 followed by the interpretation. These datasets were chosen mainly because they have been recently used by many researchers to evaluate the performance of their proposed model in predicting object-oriented software maintainability [6, 8, 9, 11, 12, 13, 16, 17, 19] and hence we wanted to be able to compare our results against this published work. The UIMS dataset contains class-level metrics data collected from 39 classes of a user interface management system, whereas the QUES dataset contains the same metrics collected from 71 classes of a quality evaluation system. Both systems were implemented in Ada. Both datasets consist of eleven class-level metrics: ten independent variables and one dependent variable. The independent variables are taken as follows:

(i) Five variables are taken from Chidambar et al. [30] : WMC, DIT, NOC, RFC, and LCOM;
(ii) Four variables are taken from Li and Henry [2, 25]: MPC, DAC, NOM, and SIZE2;
(iii) One variable is taken from traditional lines of code metric (SIZE1).
(iv) The dependent variable is a maintenance effort surrogate measure (CHANGE), which is the number of lines in the code that were changed per class during a 3-year maintenance period. A line change could be an addition or a deletion. A change in the content of a line is counted as a deletion and an addition. Table 2 defines each metric in the datasets.

TABLE 3
DESCRIPTIVE STATISTICS OF UIMS DATASET

Metric	Minimum	Maximum	Mean	Standard Deviation
WMC	0	69	11.38	15.90
DIT	0	4	2.15	0.90
NOC	0	8	0.95	2.01
RFC	2	101	23.21	20.19
LCOM	1	31	7.49	6.11
MPC	1	12	4.33	3.41
DAC	0	21	2.41	4.00
NOM	1	40	11.38	10.21
SIZE1	4	439	106.44	114.65
SIZE2	1	61	13.47	13.47
CHANGE	2	253	42.46	61.18

TABLE 4
DESCRIPTIVE STATISTICS OF QUES DATASET

Metric	Minimum	Maximum	Mean	Standard Deviation
WMC	1	83	14.96	17.06
DIT	0	4	1.92	0.53
NOC	0	0	0.00	0.00
RFC	17	156	54.44	32.62
LCOM	3	33	9.18	7.31
MPC	2	42	17.75	8.33
DAC	0	25	3.44	3.91
NOM	4	57	13.41	12.00
SIZE1	115	1009	275.58	171.60
SIZE2	4	82	18.03	15.21
CHANGE	6	217	64.23	43.13

From the descriptive statistics we noticed some observations and accordingly actions were taken. Some of them are mentioned as follows:

(i) For DIT Median and Mean value are minimum in both the system, so we draw conclusion that the use of inheritance in both systems is limited.

(ii) The values for median and mean for CHANGE (dependent variable) in the UIMS dataset is lesser than those in the QUES, which means UIMS seems to be more maintainable.

(iii) We had removed NOC from the QUES dataset because it was observed that all data points for NOC are zeros in the QUES dataset.

(iv) We had observed that the coupling between classes in QUES was higher than those in the UIMS because the medians and means values for RFC and MPC in the QUES dataset were larger than UIMS dataset.

(v) Values of Mean and Median of LCOM were almost same in both systems that mean both have almost similar cohesion.

(vi) The similar medians and means for NOM and SIZE2 in both datasets suggest that both systems had similar class sizes at the design level. However, there was a significant difference in SIZE1.

6. DISCUSSION OF RESULTS

This section consists of four subsections. In section 6.1 we have discussed experiment setup, values of various parameters initialized and values of various important parameters received after training the machine for prediction using GMDH algorithm on given pattern dataset and processing. Section 6.2 discusses the various prediction accuracy measures to compare the results of our studies with other proposed models available in literature. In section 6.3 we have summarized the values of the parameters selected in section 6.2 followed by their analysis. Section 6.4 does the interpretation of the results.

6.1 Experiment Setup and Results

In this section, the results of GMDH model were analyzed using UIMS and QUES datasets. We employed GMDH algorithm available in Neuroshell2 tool [31, 32, 33] to predict the maintainability of software. We set the parameters as shown in Table 5 while applying the proposed models on the dataset as discussed in section 5 using the tool Neuroshell2.

TABLE 5
VALUES OF PARAMETERS BEFORE EXPERIMENTAL SETUP

S.No.	Parameter	Value
1	Scale Function	[0-1]
2	GMDH type	Advanced
3	Optimization	Full
4	Maximum Variable	X1, X2, X3
5	Selection Criteria	Regularly
6	Missing value to be	Error Condition

We received following values of various parameters after we finished the process of machine learning using GMDH technique for the given data and followed by comparing the actual values with that of predicted values.

Best Formula

Where X_1, X_2,....., X_{11} are the parameters estimated by GMDH in terms of OO metrics selected in section-5 and described in Table 2 and their values are given as under:-

X_1= 2*Class-1
X_2= 2* DIT/4 -1
X_3= 2*NOC -1
X_4= 2* (MPC-2)/40 -1
X_5= 2* (RFC-17)/139 -1
X_6= 2* (LCOM -3)/30-1
X_7=2*DAC/25-1
X_8= 2* (WMC-1)/82-1
X_9= 2* (NOM-4)/53-1
X_{10}= 2* (size2-4)/78-1
X_{11}= 2* (size1-115)/894-1
Y = min (max (Change-6)/211.....),1..)

The values of various parameters and their description are given in Table 6.

TABLE 6
VALUES OF PARAMETERS CALCULATED WHEN GMDH MODEL IS APPLIED ON DATA SET

S.No	Parameter	Value	Description
1	MSE (Mean Squared Error)	0.003	It is a statistical measure of the differences between the values of outputs in the training set and the output values the network is predicting. This is mean over all patterns in file of the square of the actual value minus the predicted value, i.e., the mean of (actual - predicted)^2. The errors are squared to penalize the larger errors and to cancel the effect of the positive and negative values of the differences.
2	R-Squared	0.913	It compares accuracy of the model to the accuracy of a trivial benchmark model wherein the prediction is just the mean of all of the samples. A perfect fit would result in an R squared value of 1, a very good fit near 1, and a very poor fit less than 0.
3	Correlation Coefficient (Pearson's Linear Correlation Coefficient)	0.955	This is a statistical measure of the strength of the relationship between the actual versus predicted outputs. The r coefficient can range from -1 to +1. The closer r is to 1, the stronger the positive linear relationship, and the closer r is to -1, the stronger the negative linear relationship. When r is near 0, there is no linear relationship.
4	Normalized Mean Square Error	0.032	The Normalized Mean Square Error or Root Mean Square Error (RMSE) is a frequently used measure of the differences between values predicted by a model or an estimator and the values actually observed. It is a good measure of accuracy. The individual differences are called residuals, and it serves to aggregate differences into a single measure of predictive power.

6.2 Prediction Accuracy Measures

An important question that needs to be asked of any prediction model is "How accurate are its predictions". Based on the two values namely actual value and predicted values, researchers have stated various methods to evaluate the quality of predictions [3, 10, 20, 34, 35]. In our proposed study we evaluated and compared the OO software maintainability prediction models quantitatively with other proposed models. We used following measures:

(i). MRE (Magnitude of Relative Error): It is a normalized measure of the discrepancy between actual values and predicted values as proposed by Kitchenham in 1991 [34]. Ever since it is proposed by the author, it has become the de facto standard to measure the accuracy of software maintainability prediction. It is given as :

$$MRE = \frac{[\text{Actual Value} - \text{Predicted Value}]}{\text{Actual Value}} \quad (7)$$

(ii). MMRE (Mean Magnitude of Relative Error) : It is the mean of MRE and calculated as follows :

$$MMRE = \sum MRE \qquad (8)$$

MMRE measures the average relative discrepancy. It is equivalent to the average error relative to the actual effort in the prediction. In our study we have expressed MMRE as actual values however in some studies it is expressed in %. MMRE has been regarded as a versatile assessment criterion and has number of advantages such as it can be used to make comparisons across datasets and all kinds of prediction model types and it is independent of measuring unit and scale independent [20].

(iii). Pred : It is the measure of what proportion of the predicted values have MRE less than or equal to specified value, given by Fentom. [35]

$$Pred(q) = \frac{K}{N} \qquad (9)$$

Where q is the specified value
K is number of cases whose MRE is less than or equal to q
N is total number of cases in the datasets

In current study we have used most commonly values such as pred(0.25) and pred(0.30) in the field of software effort prediction literature so that we can compare our results.

(iv). R-Square - It is a measure of the quality of fit. It is a measure of how well the variation in the output is explained by the targets. If this number is equal to 1, then there is perfect correlation between targets and outputs [28]. It is calculated by square of the correlation coefficient. 100% R-square means perfect predictability.

(v). P-values – p-values are used for testing the hypothesis of no correlation. Each p-value is the probability of getting a correlation as large as the observed value by random chance, when the true correlation is zero. If p is small, say less than 0.05, and then the correlation i.e. R is significant [31].

6.3 Comparison with other studies

We also have compared the values of prediction accuracy measures of certain selected parameters with the studies conducted in the last decade. In Table 4 we presented the summarized performance measures of all Models studied on UIMS and QUES data set in the last decade.

TABLE 4
COMPARISON OF VARIOUS MODELS WITH REFERENCE TO THEIR PREDICTIVE PERFORMANCE FOR QUES AND UIMS SYSTEM

S No	Model Name	Reference	Max MRE	MMRE	MARE	Pred (0.25)	Pred (0.30)	Pred (0.75)	Remarks
1.	GMDH (Group Method of Data Handling) Model	Proposed in Current Study	0.983	0.210	-	0.69	0.722	0.944	Lowest MMRE Recorded in Current Study
2.	Genetic Model		0.794	0.220	-	0.66	0.722	0.972	
3.	PNN (Probabilistic Neural Networks)		0.923	0.230	-	0.68	0.75	0.944	

4.	GRNN (General Regression Neural Network) Model	[8]	4.295	0.765	-	-	-	-	
5.	ANN (Artificial Neural Network) Model	[9]	-	0.265	-	-	-	-	R- Squared 0.582 p-value 0.004
6.	Bayesian Model	[11]	1.592	0.452	-	0.391	0.430	-	
7.	Regression Tree	[11]	2.104	0.493	-	0.352	0.383	-	
8.	Backward Elimination	[11]	1.418	0.403	-	0.396	0.461	-	
9.	Stepwise Selection	[11]	1.471	0.392	-	0.422	0.500	-	
10.	MARS (Multiple adaptive regression splines)	[12]	1.91	0.32	-	0.48	0.59	-	
11.	MLR (Multiple Linear Regression)	[12]	2.03	0.42	-	0.37	0.41	-	
12.	SVM (Support Vector Machine)	[12]	2.07	0.43	-	0.34	0.46	-	
13.	ANN (Artificial Neural Network) Model	[12]	3.07	0.59	-	0.37	0.45	-	
14.	Regression Tree	[12]	4.82	0.58	-	0.41	0.45	-	
15.	TreeNets	[13]	-	0.42	-	0.58	0.65	-	
16.	Generalized Regression	[14]	-	-	0.308	-	-	-	
17.	ANFIS(Adaptive Neuro Fuzzy Inference System) Model	[14]	-	-	0.242	-	-	-	

For analyzing the results of Table 4 we have taken Max MRE (Magnitude of Relative Error) Values of other models as well as proposed models and presented them in Figure 2. It can be observed easily that MRE for GMDH is lowest which implies that it can be used as sound alternative for the prediction of maintainability. In Figure 3 a comparison has been shown between MMRE values of proposed models and other prevalent models

is presented. From Table 4 and Figure 3 it can observed that out of the thirteen algorithms selected and evaluated, the GMDH model, Genetic model and PNN model gives very competitive results and hence show their worth that they can be used in the process of software maintainability prediction. To Analyze the prediction accuracy measure we have taken values at PRED(0.25) and PRED(0.30) of all models and presented them in Figure 4. It is evident from the Table 4 and Figure 4 that prediction accuracy of GMDH network model is much better than other models. It is the only model which is close to the criterion laid by Conte *et al.* [20] and Mac Donell [3].

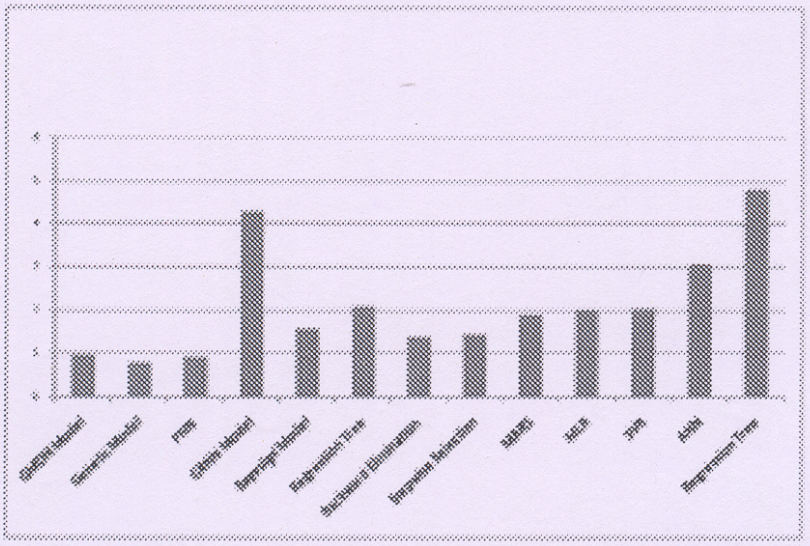


FIGURE 2: MAX MRE OF PROPOSED MODELS VERSUS OTHER MODELS

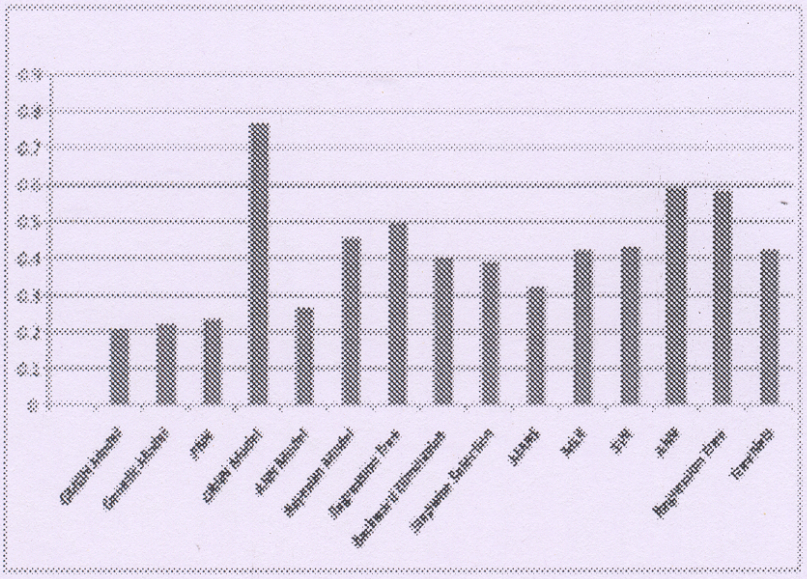


FIGURE 3: MMRE OF PROPOSED MODELS VERSUS OTHER MODELS

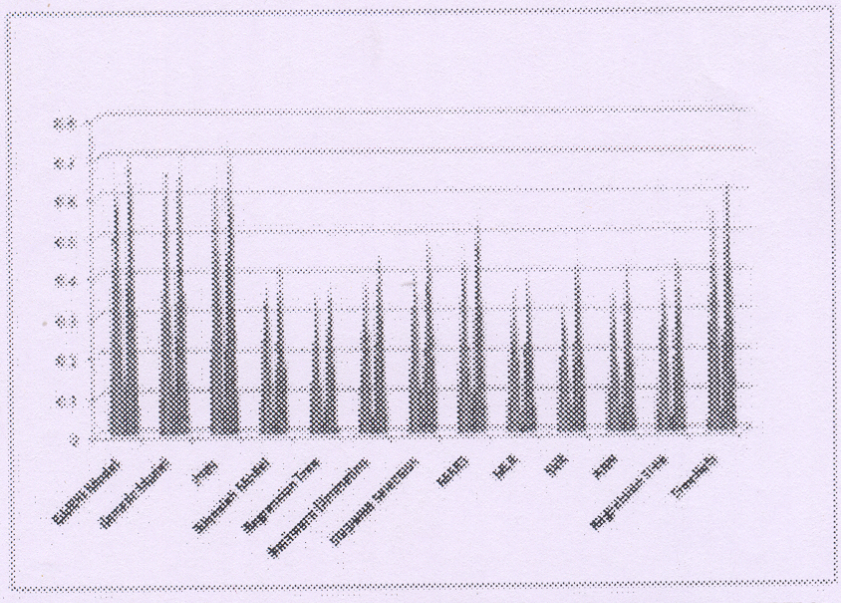


FIGURE 4: PRED(0.25) AND PRED(0.30) OF PROPOSED MODELS VS. OTHER MODELS

6.4 Analysis and Interpretation

In this paper, we have presented an empirical study that sought to build object-oriented software maintainability prediction model using following three machine learning algorithms:

(i) Group Method of Data Handling(GMDH)
(ii) Genetic Algorithm (GA)
(iii) Probabilistic Neural Network (PNN) Using Gaussian Activation function

The GMDH and GA is proposed for the first time for prediction of the software maintainability. Although Artificial Neural Network has been used previously in literature [8, 14, 37] but for the first time the Probabilistic Neural Network (PNN) along with Gaussian activation function has been applied. In this study, to draw most realistic comparison we have also analyzed the same dataset which was originally proposed by Li and Henry and earlier applied by various researchers to predict maintainability as per the details summarized in Table 4.

The criteria for prediction given by Conte *et al.* [20] and MacDonell [3] states that prediction model is considered accurate if value of pred(0.25) and pred(0.30) is greater than pred(0.75) which clearly proposed models in this study satisfies. In the literature it is also suggested that prediction accuracy of software maintenance effort prediction models is often low and thus it is very difficult to satisfy the criteria [10, 13]. It can be noticed from Table 4 that none of the prediction models satisfy the criteria. However, the GMDH model has achieved improved Pred(0.25) and Pred (0.30) over the other models in QUES and UIMS datasets, and its results are quite closer to the criteria set in literature[3, 20].

It is evident from the Table 4, the prediction accuracy of GMDH network model is much better than all the other models. At pred(0.25) its values are 0.69 which means that almost 69% predictions are less than the error of 0.25 prediction accuracy. At pred(0.30) its value is 0.722 which means that almost 72% predictions are less than the error of 0.30 prediction accuracy as compared to other models as shown in Figure 4. Following comparative analysis, it is safe to conclude that GMDH has clearly outperformed than other models. The GMDH models can predict maintainability of the OO software systems with least MMRE when compared with others models such as GRNN, ANN, Bayesians, MARS, TreeNets and SVM for QUES dataset. Hence it is clear inference that GMDH is the most accurate and best model for the predictions of software maintainability.

The SVM (Support Vector Machine) model was proposed recently by Cong *et al.* [19] for predicting maintainability using OO metrics, however it is not comparable to the current study because of the fact that their study was merely conducted on the code which was written for "Temper proof HTML web page" in C++ whereas our

study is conducted on commercially available QUES dataset written in ADA with much higher scope. Not only the sizes of the software differ to large degree but also both systems varied in to great extent with respect to their paradigm and complexity. Secondly, Max MRE and Pred(q) were not provided despite being de facto prediction standards. MARE in their model recorded as 0.218. When it is compared with current study, MMRE has been recorded better at 0.210 with GMDH model that clearly confirms higher competence even in complex environment.

7. THREATS TO VALIDITY

Like other empirical studies, limitations confronted during the current study are given as under:

(i) UIMS and QUES datasets which are considered in the study undertaken are written in ADA language. The models which have been derived in this study are likely to be valid for the code written in other object-oriented programming languages, for example C++ or Java, however, further research can only establish their usefulness in predicting the maintainability of other development paradigms.

(ii) During the process of selecting independent variables while constructing the proposed model, although utmost care has been taken and only those eleven variables are being chosen which we consider to have the strong impact on maintainability, nevertheless few other prevailing independent variables and their effect on maintainability of software also needs to be determined. Some of the widely used variables which warrant further consideration for fair comparability are Class Method Complexity (CMC), Number of Ancestor Classes (NAC), and Number of Descendent Classes (NDC), coupling through Abstract Data Type (CTA), Class Complexity (CC), Exception Handling Factor EHF, Number of Object Memory Allocation (NOMA), Average Number of Live Variables, and Average Live Variable Span etc.

(iii) Similarly, in our study although three machine learning algorithms GMDH, GA and PNN have been applied, however few others which have also gained popularity in recent times due to their effectiveness, like Random Forest, Decision Tree and Naïve Bays Network are also required to be studied on the way to empirically ascertain their merit over the proposed models.

(iv) Measuring the effectiveness of machine learning algorithms for predictions of the procedural languages is also a limitation of the proposed models.

8. CONCLUSION

Three different machine learning algorithms are used for the purpose of prediction of software maintainability. Even though many studies reported wide application of GMDH model and GA model in diverse fields for the purpose of prediction of high order input output relationship which is complex, non linear and unstructured but for the first time they are used for prediction of software maintainability. The goal of our study is to construct suitable model using machine learning algorithm for the prediction of object oriented software maintainability which are not only easy to apply but could reduce the prediction errors to minimum. The prediction performance of the machine learning algorithms based models like GMDH, GA and PNN were assessed and compared with prevailing models in terms of MMRE, Pred(0.25) and Pred(0.30). It was found that GMDH model outperformed the prevailing models as the least MMRE value is recorded. As far as Pred(0.25) and Pred(0.30) values are concerned, all the three proposed models are significantly better over others. Thus, it is concluded that GMDH network model is indeed a very useful modeling technique and it could be used as a sound alternative for the prediction of software maintainability.

As the proposed model was found suitable for estimating the software reliability in an earlier work [24], therefore with current findings it can be safely presumed that both the reliability and maintainability, which remain indispensable components of software quality, can possibly be predicted with application of same model i.e. GMDH. This will certainly reduce the challenges involved with prediction of maintainability and assists software developers to strategically utilize their resources, enhance process efficiency and optimize the associated maintenance costs.

As the current study was based on two commercial software datasets UIMS and QUES developed in ADA, therefore the authors are actively involved in carrying out further work that applies the proposed GMDH network model to more objective datasets to ascertain its authenticity for wider software paradigms. Such studies would allow us to investigate the capability of GMDH network and finally establishing a generalized model in the field of Software Quality. Further research is planned in an attempt to combine GMDH model with other data mining techniques so as to develop prediction models which can estimate the maintainability of software more accurately with least precision errors.

REFERENCES

[1]. IEEE Standard. 1219-1993 IEEE Standard for Software Maintenance. INSPEC Accession Number: 4493167 DOI: 10.1109/IEEESTD.1993.11557 Journal .1993

[2]. W. Li and S. Henry, "Object-Oriented Metrics that Predict Maintainability," *Journal of Systems and Software*, vol. 23, no 2, pp. 111-122, 1993.

[3]. S.G. Mac Donell, "Establishing relationships between specification size and software process effort in case environment," *Information and Software Technology*, vol. 39, no 1, pp. 35-45, 1997.

[4]. L Briand, C Bunse, J Daly, "A controlled experiment for evaluating quality guidelines on the maintainability of object oriented design", *IEEE Transaction on software Engineering*, vol:27, no: 6, pp 513-530, 2001, DOI: 10.1109/32.926174.

[5]. F. Fioravanti and P. Nesi, "Estimation and prediction metrics for adaptive maintenance effort of object oriented systems", *IEEE Transactions on Software Engineering*, vol. 27, no. 12, pp. 1062-1084, 2001.

[6]. KK Aggarwal, Y Singh, JK Chhabra, "An Integrated Measure of Software Maintainability", Annual *proceedings: Reliability and Maintainability Ssymposium*, pp 235-241, 2002.

[7]. S. Misra, "Modeling design/coding factors that drive maintainability of software systems" , *Software Quality Journal*, vol. 13, no. 3, pp. 297-320, 2005.

[8]. M. Thwin and T. Quah, "Application of neural networks for software quality prediction using object oriented metrics", *Journal of Systems and Software*, vol. 76, no. 2, pp. 147-156, 2005.

[9]. K.K. Aggarwal, Y. Singh, P. Chandra and M. Puri, " Measurement of Software Maintainability Using a Fuzzy Model", *Journal of Computer Sciences*, vol. 1, no.4, pp. 538-542, 2005 ISSN 1549-3636 © 2005 Science Publications.

[10]. A.D. Lucia, E. Pompella, and S. Stefanucci, "Assessing effort estimation models for corrective maintenance through empirical studies" , *Information and Software Technology*, vol. 47, no. 1, pp. 3-5, 2005.

[11]. C.V. Koten, A.R. Gray, "An application of Bayesian network for predicting object-oriented software maintainability", *Information and Software Technology Journal* , vol: 48, no : 1, pp 59-67, Jan 2006.

[12]. Y. Zhou and H. Leung, "Predicting object-oriented software maintainability using multivariate adaptive regression splines" , *Journal of Systems and Software,* vol. 80, no. 8, pp. 1349-1361, 2007.

[13]. MO. Elish and KO. Elish , "Application of TreeNet in Predicting Object-Oriented Software Maintainability: A Comparative Study", *European Conference on Software Maintenance and Reengineering,* pp 1534-5351, March 2009, DOI 10.1109/ CSMR.2009.57.

[14]. A Kaur, K Kaur, R Malhotra, "Soft Computing Approaches for Prediction of Software Maintenance Effort", *International Journal of Computer Applications*), vol 1, no. 16, *pp :* 0975 – 8887, 2010.

[15]. L Ping, "A Quantitative Approach to Software Maintainability Prediction", *International Forum on Information Technology and Applications, vol: 1, no : 1, pp : 105-108, July 2010.*

[16]. D.F. Specht , "Probabilistic Neural Networks", *Journal of Neural Networks, Elsevier,* vol. 3, no 1, pp 109-118, 1990, DOI : .org/10.1016/0893-6080(90)90049-Q.

[17]. T.D. Morco, "Controlling Software Projects: Management, Measurement and Estimation," ISBN 10: 0131717111 / 0-13-171711-1; ISBN 13: 9780131717114; Yourdon ; 1982.

[18]. M. Dagpinar and J.H. Jahnke , "Predicting Maintainability with Object-Oriented Metrics - An Empirical Comparison", *Proceedings of the 10th Working Conference on Reverse Engineering, pp 155-164, Nov* 2003.

[19]. C Jin , A.L.Jin , "Applications of Support Vector Machine and Unsupervised Learning for Predicting Maintainability using Object-Oriented Metrics", *Second International Conference on Multi Media and Information Technology ,vol 1 ,no : 1, pp 24-27, April 2010.*

[20]. S. Conte, H. Dunsmore, and V. Shen," Software Engineering Metrics and Models". Book, *Menlo Park, CA: Publisher: Benjamin-Cummings publishing co., ISBN:0-8053-2162-4*, 1986.

[21]. A. G. Ivakhnenko, "Group Method of Data Handling- A Rival of the method of Stochastic Approximation," *Soviet Automatic Control,*, vol 13, no. 3, 43-71, 1966.

[22]. A. G. Ivakhnenko and Y. U. Koppa, "Regularization of decision functions in the group method of data handling", *Soviet Automatic Control*, vol 15 no. 2, 28-37, 1970.

[23]. S.J. Farlow, "The American Statistician", vol 35, no. 4 , Nov 1981, 210-215, {The American Statistician is currently published by American Statistical Association}.

[24]. R. Mohanty, V. Ravi, M.R. Patra, "Software Reliability Prediction using Group Method of Data Handling", Proceeding of 12th International conference on *RSFDGrC, Rough Sets, Fuzzy Sets, Data Mining and Granular Computing,* LNAI 5908, pp. 344-351, 2009.

[25]. W. Li, "Another Metric Suite for Object-oriented Programming", *The Journal of System and Software*, vol 44, no : 2, pp 155–162, December 1998.

[26]. S. Muthanna, K. Kontogiannis, K. Ponnambalam, B. Stacey, " A Maintainability Model for Industrial Software Systems Using Design Level Metrics", Proceeding of Seventh Conference on Reverse Engineering, IEEE Computer Society, pp. 248, 2000.

[27]. K. Fujimoto and S. Nakabayashi, "Applying GMDH algorithm to extract rules from examples", *Systems Analysis Modeling Simulation,* vol. 43, no. 10, October 2003, pp. 1311-1319.

[28]. M.M. Ibrahiem, E.I. Emary and S. Ramakrishnan, "On the Application of Various Probabilistic Neural Networks in Solving Different Pattern Classification Problems", *World Applied Sciences Journal*, vol. 4, no. 6, pp. 772-780, 2008, ISSN 1818-4952.

[29]. S. Chidamber, R. Shyam and C. Kamerer, "Towards a metrics Suite for Object-Oriented Design Proceedings", *Proceeding of Conference on object – oriented programming systems, languages and applications, OOPSLA'91*, pp.197-211, November, 1991.

[30]. S. Chidamber and C. Kemerer, "A Metrics Suite for Object Oriented Design," *IEEE Transactions on Software Engineering*, vol. 20, no. 6, pp. pp. 476-493, June 1994.

[31]. http://www.mathworks.com.

[32]. http://www.inf.kiev.ua/GMDH-home/GMDH_abo.htm.

[33]. http://www.wardsystems.com.

[34]. B.A. Kitchenham, L.M. Pickard, S.G. MacDonell, M.J. Shepperd," What accuracy statistics really measure", *IEEE Proceedings-Software* vol. 148, no. 3, pp 81–85, 2001.

[35]. N.E. Fentom, S.L. Pfleeger, "Software Metrics: A Rigorous and Practical Approach, Second Edition", *PSW publishing Company*, 1997.

[36]. K. K. Aggarwal, Y. Singh, A. Kaur and R. Malhotra, "Application of Artificial Neural Network for Predicting Maintainability using Object-Oriented Metrics", *World Academy of Science*, pp. 140-144, 2006

[37]. KK Aggarwal, Y Singh, A Kaur, R Malhotra, "Application of artificial neural network for predicting fault proneness models" , *International Conference on Information Systems, Technology and Management (ICISTM 2007)*, pp.12-13, March 2007.

[38]. KK Aggarwal, Y Singh, A Kaur, R Malhotra, "Analysis of object-oriented metrics", International Workshop on Software Measurement *(IWSM)*, 2005.

ABOUT THE AUTHORS

Ruchika Malhotra is an Assistant professor with Department of Software Engineering, Delhi Technological University (formerly Delhi College of Engineering), India. She was an Assistant professor with University School of Information Technology, Guru Gobind Singh Indraprastha University, India. She is the Executive Editor of Software Engineering International Journal, Delhi Technological University. She received her doctorate and masters degree from the University School of Information Technology, Guru Gobind Singh Indraprastha University, India. She is a co-author of book titled "Object Oriented Software Engineering" published by PHI Learning. Her research interests are in improving software quality, statistical and adaptive prediction models for software metrics, neural nets modeling, and the definition and validation of software metrics. She has 60 publications in international journals and conferences. Her paper titled Analysis of object oriented Metrics was published as a chapter in the book Innovations in Software Measurement (Shaker -Verlag, Aachen 2005). She can be contacted by e-mail at ruchikamalhotra2004@yahoo.com.

Anuradha Chug is Assistant Professor in the Department of University School of Information and Communication Technology (USICT), Guru Gobind Singh Indraprastha University, New Delhi. She has long teaching experience of almost 19 years to her credit as faculty and in administration at various educational institutions in India. She has worked as guest faculty in Netaji Subhash Institute of Information and Technology, Dwarka, New Delhi and Regular Faculty at Government Engineering College, Bikaner. She has also worked as Academic Head, Aptech, Meerut and Program Coordinator at Regional Centre, IGNOU, Meerut. Her areas of research interest are Software Engineering, Neural Networks, Analysis of Algorithms, Data Structures and Computer Networks. She is currently pursuing her Doctorate degree from Delhi Technological University. She has earlier achieved top rank in her M.Tech (IT) degree and conferred the University Gold Medal in 2006 from Guru Gobind Singh Indraprastha University. Previously she has acquired her Master's degree in Computer Science from Banasthali Vidyapith, Rajasthan in the year 1993. She has also presented number of research papers in national/ international seminars, conferences and research journals. She can be contacted by E-mail at a_chug@yahoo.co.in

THERMODYNAMIC ANALYSIS FOR IMPROVEMENT IN THERMAL PERFORMANCE OF A SIMPLE GAS TURBINE CYCLE THROUGH RETROFITTING TECHNIQUES (INLET AIR EVAPORATIVE COOLING, STEAM INJECTION & COMBINED IAC AND STIG)

Shyam Agarwal*[1] and R.S. Mishra[1]
[1]Department of Mechanical Engineering, Delhi College of Engineering, Bawana Road, Delhi-110042, India
* Corresponding author – Tel.: 91+9013353987; E-mail address: sh_agar@rediff.com

Abstract
Retrofitting technologies [inlet evaporative cooling system and steam injected gas turbine] have been applied on simple gas turbine cycle for performance improvement followed by parametric analysis. The performance improvement has been thermodynamically analysed and discussed for retrofitted techniques followed by performance studies. The parametric study predicts that retrofitting techniques (FCS and STIG) improves net power output, thermal efficiency, power generation efficiency , first law efficiency and exergetic efficiency (second law efficiency) while heat rate falls with a considerable increment in fuel consumption. Exergy analysis showed that combustion chamber and turbine are most sensitive components of retrofitted system. The results show that the power output , thermal efficiency , exergetic efficiency and fuel-air ratio have been enhanced 3.1% , 0.18%, 0.2% and 1.0% respectively while heat rate falls 0.6% by FCS technology. The power output , thermal efficiency , exergetic efficiency and fuel-air ratio have been improved 27.4% , 3.5%, 25.8% and 14.4% respectively while heat rate falls 10.2% by STIG technology. The analysis shows that STIG technology is better than FCS and the combined FCS & STIG technology enhance the power output , thermal efficiency , exergetic efficiency and fuel-air ratio 30.5% , 3.5%, 25.7% and 15.2% respectively and reduces the heat rate 10.4%.
Keywords: *Gas turbine, Retrofitting, FCS, STIG, Exergy*

Nomenclature

AP	Approach point (^{0}C)		
PP	Pinch point(^{0}C)	**Superscript**	
TIT	Turbine Inlet Temperature (K)	,	fraction of gas phase at dead state
wbt	Wet bulb temperature(^{0}C)	1	fraction of gas phase before combustion
M	Molecular weight (kg/kmol)	2	fraction of gas phase after combustion
$\dot{E}$	Exergy rate (kW)	CH	chemical
U	Internal energy (kJ)	PH	Physical
$\dot{m}$	mass flow rate (kg/s)	PT	potential
Subscripts		**Greek symbol**	
sup	superheated	η	efficiency
Th	thermal	λ	fuel-air ratio
HRSG	Heat recovery steam generator	ω	steam-injection ratio
GEN	Heat generator	€	Exergetic efficiency
f	fuel	**Acronyms and abbreviations**	
REG	Regenerator	FCS	Fog cooling system
dbt	Dry bulb temperature(^{0}C)	STIG	Steam injection gas turbine

1. Introduction
Simple gas turbine power generation systems are widely used in Indian industries due to quick startup and shutdown capabilities. These units are mostly used to fulfil the peak load demand but unfortunately suffer from very low overall efficiency and reduction in power output during the summer season, when electricity demand is the highest. To investigate the anticipated power shortage, retrofitting projects have been seriously gaining momentum to covert these existing simple gas turbine systems into relatively advanced cycle units resulting in both improved efficiency and power output.

From the fundamnetal thermodynamic point of view, the reason for low thermal efficiency of simple gas turbine are high back work ratio (large part of turbine work used to compress the inlet air) and substantial amount of available energy loss due to high temperature exhaust (ofter above 500^0C) caused by shaft rotation of turbine at a relativelely high back pressure. This waste heat recovery from the gas turbine exhaust can be utilized to improve generation efficiency through various modifications to basic cycle such as gas to gas recuperation, steam injection, chemical recuperation, inlet air cooling and combined cycle etc. Among many well-established technologies, the combined cycle is rated as the most efficient way to recover the energy from the exhaust for boosting the capacity and efficiency of power generation. However, the combined cycle technology is unsuitable for daily on-off operation pattern units due to low mobility (start-up time).

Recently, the steam injection gas turbine (STIG) and inlet air cooling (IAC) by evaporation are the most common practices to enhance the performance of power generation. Both feature can be implemented in the existing basic system without major modification to the original system integration in a economic way. In the STIG technology, the steam generated from the heat-recovery generator (HRSG) is injected into the combustion chamber. Both steam from the HRSG and air from the compressor receive fuel energy in the combustion chamber and both simultaneously expand inside the existing turbine to boost the power output of turbine. It should be noted that the required injected steam pressure is obtained from a pump fitted in the steam unit. The net power output produced by steam is significantly higher than that of air in terms of per unit mass flow rate, due to very small pumping work compared to compressor. The thermodynamic processes of a simple cycle gas turbine can be approximately modeled as a Brayton cycle, in which the back work ratio is usually very high, and the exhaust temperature is often above 500 ^{0}C. A high exhaust temperature implies there is plenty of useful energy rejected to the environment. The recovery of this wasted energy can otherwise be used to improve either the power generation capacity or efficiency via retrofitting to the basic cycle such as STIG and inlet air cooling by evaporation technology. Since, the specific heat of steam and hence enthalpy is relatively higher than air at a certain temperature, the STIG method is a very effective alternative to increase the efficiency and boost the power output of gas turbine. Kumar et al.[1] have been developed design methodology for parametric study and thermodynamic performance evaluation of a gas turbine cogeneration system (CGTS). Wang & Chiou [2] suggested that application of IAC and STIG technique can boost the output and generation efficiency. They concluded that implementing both STIG and IAC features cause more than a 70% boost in power and 20.4 % improvement in heat rate. Bouam et al.[3] have studied combustion chamber steam injection for gas turbine performance improvement during high ambient temperature operations. Their research study is to improve the principal characteristics of gas turbine used under extreme ambient condition in Algerian Sahara by injecting steam in the combustion chamber. Srinivas et al.[4] have worked out on sensitivity analysis of STIG based combined cycle with dual pressure HRSG. They concluded that steam injection decreases combustion chamber and gas reheater exergetic loss from 38.5 to 37.4% compared to the case without steam injection in the combustion chamber. Minciuc et al.[5] have presented thermodynamic analysis of tri-generation with absorption chilling machine. They have focused on solutions of tri-generation plants based on gas turbine or internal combustion engine with absorption chilling machine. Moran [6] has developed design and economic methodology for the gas turbine cogeneration system. Nishida et al.[7] have analyzed the performance characteristics of two configuration of regenerative steam-injection gas turbine (RSTIG) systems. They concluded that the thermal efficiencies of the RSTIG systems are higher than those of regenerative, water injected and STIG systems and the specific power is larger than that of regenerative cycle.

The IAC technology is simply to cool down the inlet air entering the compressor with a cooler. Due to this, the compressor consumes less work and can compress more air per cycle to increase the capacity of the gas turbine. Different IAC options are evaporative cooling, mechanical chiller, absorption chiller and thermal energy storage, etc. applied in gas turbine power augmentation. Among them evaporative cooling is the cheapest one. Sinha & Bansode[8] have studied the effect of fog cooling system for inlet air cooling. They concluded that performance parameters indicative of inlet fogging effects have a definite correlation with the climate condition (humidity and temperature) and showed improvement in turbine power and heat rate with inlet fogging. Chaker et al.[9] have developed the formulation for fog droplet sizing analysis and discussed various nozzle types, measurement and testing. This study describes the different available measurement techniques, design aspects of nozzles, with experimental and provides recommendations for a standardized nozzle testing method for gas turbine inlet air fogging. Salvi & Pierpaloli[10] have studied optimization of inlet air cooling systems for steam injected gas turbines. They proposed the technique of compression inlet air cooling through an ejection system supplied by the exhaust heat of the gas turbine. Bassily[11] has studied the performance improvements of the intercooled, reheat and recuperated gas turbine cycle using absorption inlet-cooling and evaporative after-cooling. A parametric study of the effect of pressure ratio, ambient temperature, ambient relative-humidity, turbine's inlet-temperature (TIT), and the effectiveness of the recuperated heat-exchanger on the performance of varieties of cycles is carried out. Bhargava & Homji[12] have studied parametric analysis of existing gas turbines with inlet

evaporative and overspray fogging. This study shows the effects of inlet fogging on a large number of commercially available gas turbines.

Although many efforts have been focused to apply either STIG technology or the IAC method to enhance the gas turbine's performance, very limited studies are available regarding simultaneous integration of STIG and IAC for the same system. In this study, a simple cycle generation unit is carried as base data and STIG and IAC features are retrofitting by added in a sequencing of the system. The benefits obtained from either the STIG or IAC can be distinguished or the integration effect from the combined STIG and IAC can be realized. The performance analysis has been carried out by varying ambient temperature, humidity ratio and steam injection ratio.

System Description

The simple cycle gas turbine system integrated with both IAC and STIG featuring are shown in Figure 1. The basic unit includes compressor, combustor, gas turbine and a generator. A heat recovery steam generator (HRSG) was installed at the downstream exit of the turbine (state point 5) in order to recover the heat from the exhaust gases. The fraction of superheated steam generated from the HRSG is used for STIG (state point 9) and the remaining steam is used for process application. An evaporative fog cooling system (FCS) is installed to cool the ambient air (state point $1'$) as shown in Fig. 1. Fog cooling is an active system which uses very fine fog droplets of high pressure water injected through special atomizing nozzles located at discrete points across the inlet duct at high pressure to create the cooling effect. The amount of fog is to be monitored base on dry and wet bulb ambient conditions to achieve the required cooling. A typical fog cooling system consists of a high pressure pump skid connected for feeding to an array of manifolds located at a suitable place across the compressor inlet duct. The manifolds have a requisite number of fog nozzles[6] which inject very fine droplets of water into the inlet air. The discharge through each nozzle is around 3ml/s and produces 3 billion droplets per second. The fine fog evaporates very fast, thus dropping inlet air temperature.

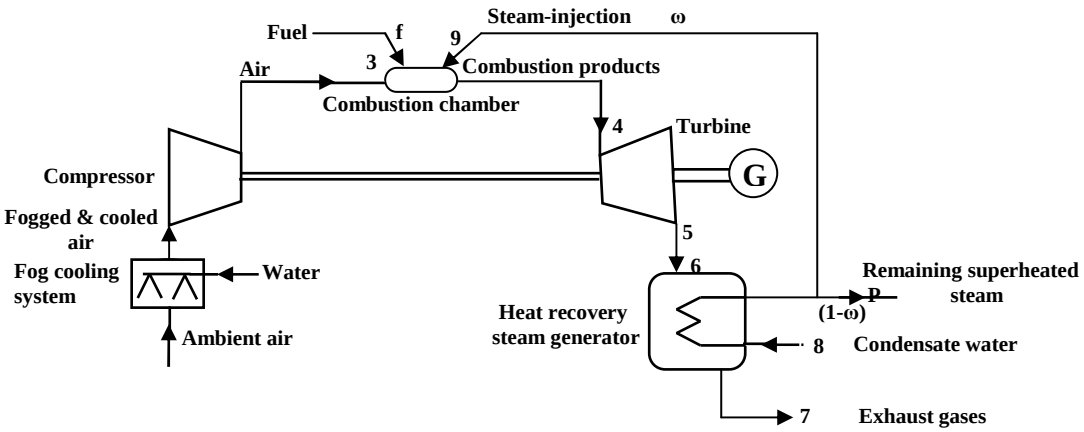


Fig. 1- Simple cycle gas turbine with fog cooling

Modeling and Computer simulation

Formulation

The following assumptions have been considered for the present study:

1. The composition of dry air has been assumed in terms of molar fraction of 1mole of air is: (N_2 = 0.78981, O_2 = 0.20989, CO_2 = 0.00031 and H_2O = 0).
2. The heat loss from the combustion chamber is 2% of the fuel lower heating value. All other components operate without heat loss.
3. Fog cooling system has been maintained for 100% saturation of ambient air at wet bulb temperature of air.
3. The pressure of water injected from the nozzle into the evaporative cooling chamber has been assumed 138 bar and converts into the fog (fine droplets), absorbs latent heat of air through adiabatic mixing.
4. Combustion chamber has been maintained at constant temperature in the presence of steam (in case of steam injection).

A computer program has been developed to formulate and simulate the retrofitting techniques over simple gas turbine with a set of steady-state governing equations including mass, energy, entropy and exergy balances using control volume analysis sequentially for compressor, combustor, gas turbine and HRSG.

For complete combustion of natural gas (methane) with steam injection in the combustion chamber, chemical equation takes the following form:

$$\lambda CH_4 + \left[X_{1,N2}N_2 + X_{1,O2}O_2 + X_{1,CO2}CO_2 + X_{1,H2O}H_2O\right] + wH_2O$$

$$\rightarrow \left[1+\lambda+w\right] \quad \left[X_{2',N2}N_2 + X_{2',O2}O_2 + X_{2',CO2}CO_2 + X_{2',H2O}H_2O\right] \qquad \ldots \quad (1)$$

Mole fraction of N_2 $$X_{2,N2} = \frac{X_{1,N2}}{1+\lambda+w} \quad (2)$$

Mole fraction of O_2 $$X_{2,O2} = \frac{X_{1,O2} - 2\lambda}{1+\lambda+w} \quad (3)$$

Mole fraction of CO_2 $$X_{2,CO2} = \frac{X_{1,CO2} + \lambda}{1+\lambda+w} \quad (4)$$

Mole fraction of H_2O $$X_{2,H2O} = \frac{X_{1,H2O} + 2\lambda + w}{1+\lambda+w} \quad (5)$$

where ω is the steam injection ratio defined as the ratio of mass of steam injected to the mass of air supplied.

$$\omega = \dot{m}_s/\dot{m}_a\,,\ \omega' = \dot{m}_s/\dot{m}_g\,,\ \omega' = \omega/(1+\lambda),\ \omega'' = \frac{\dot{m}_s}{\dot{m}_f},\quad \omega'' = \frac{\omega}{\lambda} \quad (6)$$

where ω is the mass of steam injected to the mass of air supplied, ω' is the ratio of mass of steam injected to the mass of combustion gases formed and ω'' is the ratio of mass of steam injected to the mass of fuel supplied[2]. The maximum amount of permitted STIG is 20% of mass flow rate of inlet air[2].

The heat transfer between exhaust gases and condensate water has been taken place in water heat recovery boiler where superheated steam is generated.

$$m_{exh}(h_6 - h_7) = m_w(h_{sup} - h_{cond}) \quad (7)$$

where m_{exh} and m_w are mass flow rate of exhaust gases of turbine and condensate water, h_6, h_7, h_{sup} and h_{cond} are enthalpies of exhaust gases at state 6 and 7, super-heated steam and condensate water.

$$T_{PP} = T_{sat} + PP \quad (8)$$

$$T_{AP} = T_{sat} - AP \quad (9)$$

where T_{pp} , T_{sat} and T_{AP} are Pinch point temperature, saturation temperature of water and approach point temperature. PP is the pinch-point difference and AP is the approach point difference from saturation temperature.

The temperature of air after fog cooling can be obtained from an energy balance on the dry air, water spray and air-born water vapour before and after the system. Assuming adiabatic mixing, the energy gained by the sprayed water is balanced by the energy lost by the dry air, and the original air-born mixture, after cooling such that:

$$m_w(h_{v1} - h_{w1'}) = m_a(h_{a1'} - h_{a1}) + \omega_1 m_a(h_{v1'} - h_{v1}) \quad (10)$$

where m_w and $h_{w1'}$ are the mass flow rate and enthalpy of cooling water , m_a is the mass flow rate of dry air, ($h_{a1'} - h_{a1}$) is the enthalpy change of dry air, ($h_{v1'} - h_{v1}$) is the enthalpy change of water vapour during cooling.

The humidity ratio ($\omega_{1'}$) can be specified as:

$$\omega_{1'} = \frac{0.622\ P_{v1'}}{P_{1'} - P_{v1'}} \quad (11)$$

where $P_{v1'}$ is the partial pressure of water vapour and $P_{1'}$ is the total atmospheric pressure. From conservation of mass, the amount of water evaporated is equal to the mass of water vapour at point 1minus the water vapour originally in the air at point $1'$.

$$\Delta m_w = (\omega_1 - \omega_{1'}) m_a \quad (12)$$

where ω_1 is the humidity ratio of air after cooling, The partial pressure of water vapour can be found from the respective relative humidity (ϕ) by:

$$P_v = \phi P_{sat} \quad (13)$$

where P_{sat} and P_{sat} are the saturation pressures of water for the corresponding temperature. Pressure loss in the adiabatic mixing is neglected.

The enthalpy, entropy, and exergy can be determined at each state point using mass and energy balances. The performance parameters required for the thermodynamic analysis of retrofitted systems are given below:

Thermal Efficiency (η_{Th}): Thermal efficiency of a thermal system is defined as the ratio of net work output ($\dot{W}_{net}$) to the total heat input ($\dot{Q}_f$) of the fuel.

$$\eta_{Th} = \frac{\dot{W}_{net}}{\dot{Q}_f} \quad (14)$$

Generation Efficiency (η_{Gen}): Generation efficiency of a thermal system is defined as the ratio of electrical power output (W_{el}) to the total heat input of the fuel(Q_f).

$$\eta_{Gen} = \frac{W_{el}}{Q_f} \quad (15)$$

Heat-Rate (HR in kJ/s/kW): Heat rate is defined as the ratio of heat produced by the fuel ($\dot{Q}_f$) to the electrical power output ($\dot{W}_{el}$) of the thermal system.

$$HR = \frac{\dot{Q}_f}{\dot{W}_{el}} \quad (16)$$

Specific Fuel-Consumption (SFC): Specific fuel–consumption of a thermal system is defined as the ratio of mass of fuel to the net work output. It is reciprocal of specific net work (W_{spec}).

$$SFC = \frac{\dot{m}_f}{\dot{W}_{net}} \quad (17)$$

First–Law Efficiency (η_I): The ratio of all the useful energy extracted from the system (electricity and process heat) to the energy of fuel input is known as first-law efficiency. First-law efficiency is also known as fuel utilization efficiency or utilization factor or energetic efficiency. By definition,

$$\eta_I = \frac{\left(\dot{W}_{el} + \dot{Q}_{Pro}\right)}{\dot{Q}_f} \quad (18)$$

where $\dot{Q}_{Pro}$ is process heat rate.

Second–Law Efficiency (η_{II}): Since exergy is more valuable than energy according to the second law of thermodynamics, it is useful to consider both output and input in terms of exergy. The amount of exergy supplied in the product to the amount of exergy associated with the fuel is a more accurate measure of thermodynamic performance of a system, which is called second-law efficiency. It is also called exergetic efficiency (effectiveness or rational efficiency). By definition,

$$\eta_{II} = \frac{\left(\dot{W}_{el} + \dot{E}_{pro}\right)}{\dot{E}_f} \quad (19)$$

where $\dot{E}_{Pro}$ is the exergy content of process heat and $\dot{E}_f$ is the exergy content of fuel input.

Exergy-Destruction Rate ($\dot{E}_{DR}$)-The component exergy destruction rate can be compared to the total exergy destruction rate within the system.

$$\dot{E}_{DR} = \frac{\dot{E}_D}{\dot{E}_{D,tot}} \quad (20)$$

Results and Discussion

In the present study following configurations with retrofitting have been studied in comparison to simple gas turbine cycle:

(i) Simple gas turbine cycle with inlet air cooling (IAC)

(ii) Simple gas turbine cycle with STIG
(iii)Simple gas turbine cycle with both IAC and STIG.

The initial conditions of these system analysis are as shown in Table 1. In the calculation the steady state operation is investigated without considering the turbine blade cooling. The performance analysis of these retrofitted gas turbine system is done by preparing a computer program in EES and validated with Moran[6]. The temperature, pressure and gas concentration in each component are calculated by taking into consideration of the compositions and proportions of gases and consequently various parameters (refer Table 1) and exergy loss of these systems are estimated. The net power output and power generation efficiency are 30MW and 29.93% (shown in Table 2) respectively. Attachment of evaporative coolers with simple cycle slightly improves the performance parameters. Due to reduction in compressor work the net power output increases by 3%. The impact of evaporative cooling will be higher in dry summer season where dry bulb temperature is higher and R.H is lower.

Simple gas turbine cycle with STIG (for steam injection ratio 0.1316) significantly improves the system efficiencies. The net power output increases by 27.4%. Combination of simple cycle with STIG and evaporative cooling further improves the system performance. It is quite obvious that contribution of STIG is quite significant.

1.Comparison of simple cycle gas turbine with and without FCS shows that net power output increases by 3.1% and various efficiencies increase by 0.6% while heat rate decreases by 0.6% using FCS technology.

2.Comparison of simple cycle gas turbine with and without STIG shows that net power output increases by 27.4% and thermal efficiency increases by 11.4% while heat rate decreases by 10.2% using STIG technology. The power output, power generation efficiency improve appreciably while First-law efficiency and process heat fall with increasing amount of STIG.

3.Comparison of simple cycle gas turbine with and without FCS and STIG shows that net power output increases by 30.5% and thermal efficiency increases by 11.59% while heat rate decreases by 10.4% using FCS and STIG combine technology. The generation efficiency and net power output increases while First-law efficiency (utilization factor) and process heat decreases with increasing amount of STIG.

Table 1: Essential input parameters for simple gas turbine cycle and retrofitted systems

Parameters	**Simple gas turbine cycle+Fog cooling+STIG**
Ambient air temperature at state 1′,in K	298.15
Ambient air pressure at state1′, in bar	1.013
Ambient air relative humidity at state1′, in %	60
Spray water temperature at state1′, in K	298.15
Spray water pressure at state1′, in bar	138
Air inlet pressure to compressor (P_1), in bar	1.013
Air inlet temperature to compressor, (T_1) in K	298.15
Relative humidity of inlet air to compressor at 1, in %	100
Pressure ratio of compressor (r_p)	10:1
Isentropic efficiency of compressor (η_{SC}), in %	0.86
Isentropic efficiency of Turbine (η_{ST}), in %	0.86
Lower heating value of fuel (LHV), in kJ/kmol	802361
Mass flow rate of air (m_a), in (kg/s)	81.4
Turbine inlet temperature (TIT) or maximum cycle temperature (T_4), in K	1520
Injection pressure of fuel (methane) (P_f), in bar	12
Injection temperature of fuel (methane) (T_f), in K	298.15
Pressure drop in combustion chamber , in %	5
Exhaust pressure of combustion products after HRSG (P_7), in bar	1.013
Exhaust temperature of combustion products after HRSG (T_7) in K	403.15
Pressure of steam generation (P_9) in bar	20
Pressure of condensate water at inlet of HRSG (P_8), in bar	20
Temperature of condensate water at inlet to HRSG (T_8), in K	298.15
Pressure drop in HRSG on the gas side, in %	5
Amount of steam injected , in (% of the mass flow rate of the air)	10
Temperature of superheated steam STIG (T_9), in K	753.15
Approach point , in K	2

Pinch point, in K	20

TABLE 2: Comparison of various performance parameters of simple gas turbine cycle and retrofitted cycles

Performance Parameter	Simple gas turbine cycle	Simple gas turbine cycle with Fog cooling	Simple gas turbine cycle with STIG	Simple gas turbine cycle with Fog cooling & STIG
First law efficiency (%)	30.54	30.72	72.57	72.69
Second law efficiency (%)	29.51	29.70	55.3	55.2
Power generation efficiency (%)	29.93	30.11	33.33	33.4
Thermal efficiency (%)	30.54	30.72	34.01	34.08
Fuel-air ratio	0.0431	0.04355	0.0493	0.04967
Steam injection ratio (per kg of mass of fuel)	---	---	0.1316	0.13
Heat rate (kW/kWh)	12029	11958	10800	10780
Specific net work (kJ/kg of fuel) or specific power (kW/kg/s of fuel)	15274	15364	17012	17043
Specific fuel consumption(kg/kWh)	0.2357	0.2343	0.2116	0.2112
Work-heat ratio (kJ/kJ)	---	---	0.8823	0.8826
Power-to-heat ratio (kW/kJ/s)	---	---	0.8647	0.8649
Specific work ISO (kW-s/kg of air)	361.2	367.2	460.2	464.8
Turbine work (MW)	56.48	57.31	64.71	65.54
Compressor work (MW)	26.48	26.38	26.48	26.39
Net Power output (MW)	30	30.93	38.23	39.15
Electric work done (MW)	29.4	30.31	37.46	38.36
Process heat (MW)	---	---	43.32	44.35

Table 3 represents the comparison of exergy destruction rate in the components of simple cycle and retrofitted cycles. The exergy destruction rate represents the waste of available energy. Exergy destruction of all components have been calculated to enhance the understanding of cycle performance. Table 3 presents the exergy destruction of each cycle component after retrofitting. While examine the exergy destruction for all components, the combustor has the largest exergy destruction and shows the major location of thermodynamic inefficiency because of large irreversibility arising from the combustion reaction and heat transfer. Steam injection will increase the exergy destruction due to more mixing and combustion in the combustor. The exergy-losses at position 7 (see Fig.1) is considered as exergy loss through stack. Since part of exhaust heat is recovered in HRSG, the exhaust exergy out of stack can be reduced substantially after retrofitting. The exergy losses through stack will not only waste the available exergy but also dump the thermal pollution to our living environment.

Exergetic efficiency for each component can be defined as the ratio of $\dot{E}_R$ to $\dot{E}_S$. Where $\dot{E}_S$ is the exergy rate supplied to the component and $\dot{E}_R$ is the exergy rate recovered from the component. For a retrofitted cycle with fog cooling and STIG, exergetic efficiencies of compressor, turbine, combustor and HRSG are respectively 91%, 93%, 68% and 75%, among which gas turbine and compressor have a higher exergetic efficiency. This implies most of the exergy destruction in compressor and combustor are inexistable. It is interesting to note that although the exergy destruction rate of combustor is the highest, exergy efficiency of the combustor is higher than that of HRSG. Therefore, there is a greater improvement margin exists for HRSG than for combustor.

Table 3: Comparison of exergy destruction in the components for simple gas turbine cycle and retrofitted cycles

Components	Simple gas turbine cycle	Simple gas turbine cycle+Fog cooling	Simple gas turbine cycle+STIG	Simple gas turbine cycle+Fog cooling+STIG

	Exergy destruction rate (MW)	**€**	**Exergy destruction rate (%)**	**Exergy destruction rate (MW)**	**€**	**Exergy destruction rate (%)**	**Exergy destruction rate (MW)**	**€**	**Exergy destruction rate (%)**	**Exergy destruction rate (MW)**	**€**	**Exergy destruction rate (%)**
Combustion chamber	33.31	0.67	46.47	34.22	0.67	46.42	39.633	0.69	63.89	40.66	0.68	63.72
Gas turbine	3.193	0.95	4.16	3.336	0.95	4.53	5.018	0.93	8.09	5.075	0.93	7.95
Air compressor or Air compressor assembly	1.892	0.93	2.47	2.366	0.91	3.2	1.892	0.93	3.05	2.367	0.91	3.71
HRSG	---	---	---	---	---	---	13.114	0.75	21.14	13.283	0.75	20.82
Stack-loss	33.28	---	46.43	33.8	---	45.85	2.379	---	3.83	2.421	---	3.79
Overall plant ($\sum \dot{E}_D$) Power output (MW)	71.675 30	---	100	73.722 30.93	---	100	62.036 38.23	---	100	63.806 39.15	---	100
Total exergy destruction per MW of output($\sum \dot{E}_D / MW$)	2.39	---	---	2.38	---	---	1.623	---	---	1.629	---	---

Percentage exergy destruction rate is the exergy destruction rate within a component as a percentage of the total exergy destruction rate within the simple gas turbine system.

1.Comparison of simple cycle gas turbine with and without FCS shows that exergetic efficiency gets also improve by 0.64%. However fuel-air ratio increases by 1%. The exergy destruction gets reduced into compressor and turbine due to low inlet temperature of air.

2.Comparison of simple cycle gas turbine with and without STIG shows that exergetic efficiency gets also improve by 87.4% however fuel-air ratio increases by 14.4%. Exergy destruction increases in each system component except air compressor due to mixing of steam and air. The exergetic efficiency improves appreciably exergy destruction rate (%) of combustion chamber falls with increasing amount of STIG.

3.Comparison of simple cycle gas turbine with and without FCS and STIG shows that exergetic efficiency gets also improve by 87.06% however fuel-air ratio increases by 15.2%. The exergy destruction gets increased in each system component due increasing mass flow rate. Exergy destruction rate (%) of system component does not show much variation with relative humidity however with increasing amount of STIG, exergy destruction rate of each component increases except combustion chamber and compressor.

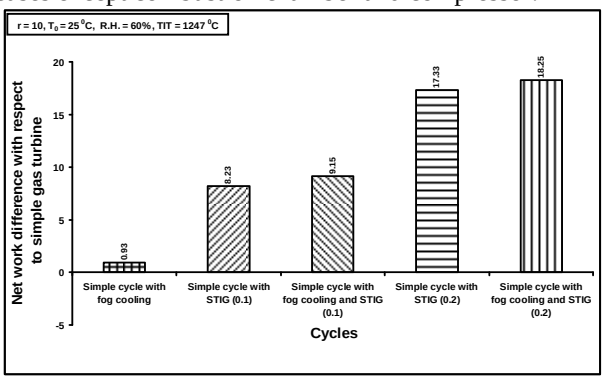


Fig. 2- Comparison of net work output for different cycles

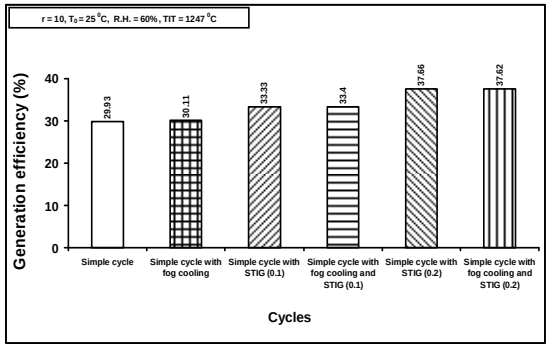


Fig. 3- Comparison of Generation efficiency for different cycles

The trend of graphs shows that the combination of fogging and STIG with simple cycle gas turbine cycle is a good approach to enhance the performance of the system on the basis of first law as well as second law. Fig.2 & 3 predict that power output is maximum in case of simple cycle with fogging and STIG and generation efficiency is maximum for STIG only. The power output increases as the mass flow rate increases and generation efficiency reduces minutely due to higher fuel-air-ratio.

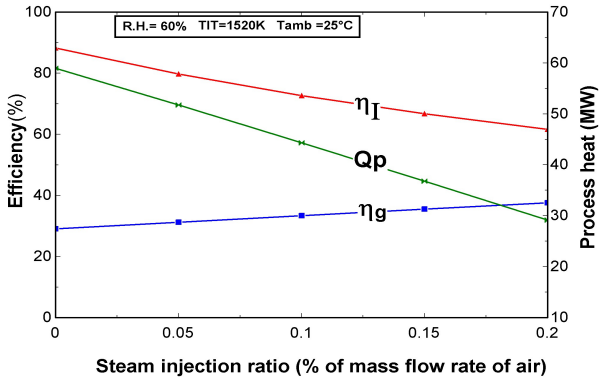


Fig. 4- The effect of steam injection ratio on First-law efficiency, generation efficiency and process heat

The Fig.4 shows the effect of STIG (0-0.2% of mass flow rate of air) on generation efficiency, first law efficiency and process heat for fixed inlet air conditions as the air gets saturated up to 100% R.H. The first law efficiency falls with the increasing amount of steam injection ratio. The reason for decrease in First-law efficiency is that the slope of process heat is sharper than the slope of generation efficiency. The generation efficiency increases while the process heat falls with increasing amount of steam injection ratio along with the fogging of air up to 100% R.H. The graph predicts that steam injection effects the generation ef ficiency, first-law efficiency and process heat more than the fog cooling of inlet.

The exergy destruction rate (MW) for different system components with different amount of STIG has been shown in Fig.5. The power output increases for large amount of STIG due to increasing mass flow rate of air. While the exergy destruction rate increases into the Combustion chamber, turbine and HRSG. The exergy destruction in combustion chamber is highest among all the system components due to highest temperature of combustion chamber. The graph predicts that steam injection increases the exergy destruction in combustion chamber due to mixing of high temperature superheated steam to the combustor highers the overall temperature.

The exergy destruction rate (MW) per MW of power output for different system components with different amount of STIG has been shown in Fig.6. Due to significant increase in power output the rate of exergy destruction (MW) per MW of power output reduces for combustion chamber, compressor, HRSG and Stack - gases while increases for gas turbine due to increasing mass flow rate with steam of lower exergy.

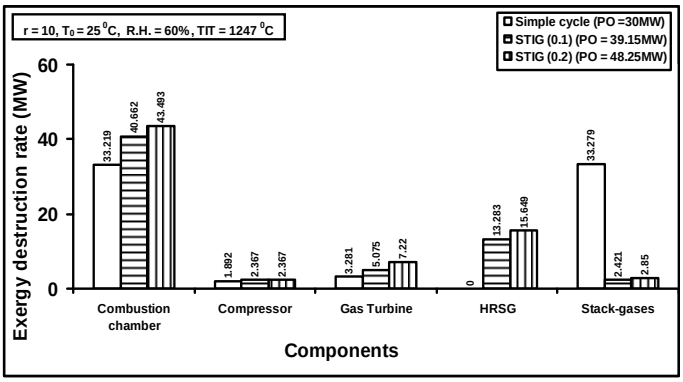


Fig. 5- Comparison of Exergy destruction rate of system components per MW of output for retrofitted cycle (Fog cooling and STIG) with simple cycle.

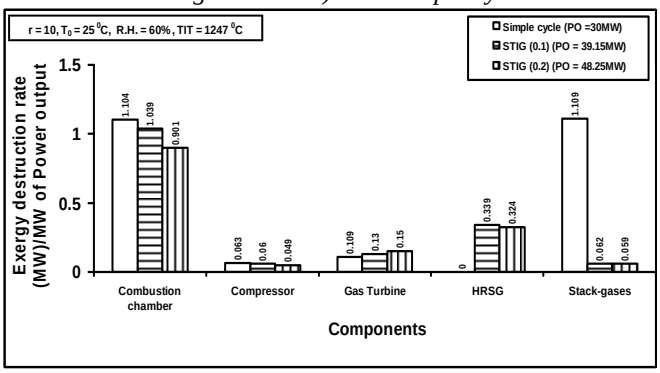


Fig. 6- Comparison of Exergy destruction rate (MW) of system components for retrofitted cycle (Fog cooling and STIG) with simple cycle.

2. Conclusions

In this study, the performance characteristics of simple gas turbine retrofitted with either fog cooling system (FCS) or /and steam injected gas turbine (STIG) technology have been analyzed and the various efficiencies and net power output have been investigated, In addition, various performance parameters have also been compared on the basis of first law as well as second law (exergy) analysis.

- During summer season as temperature increases, the density of ambient air decreases and hence the mass flow rate decreases which consequences low power output. The FCS technology reduces the ambient temperature and increases relative humidity which ultimately increases mass flow rate and hence power output increases. In this study, the power output improves by 3.1% and thermal efficiency by 0.6% using FCS technology retrofitted with simple cycle gas turbine.
- As the exhaust gases leaving from the turbine have very high temperature and fundamental thermodynamic analysis predicted that low efficiency of gas turbine has been resulted. The methods used in the past to improve the efficiency were focused on either increasing the expansion work or decreasing the compression work. In this study, the modification of STIG and FCS to simple gas turbine enhanced the thermal efficiency up to 11.59% and the power output improved by 30.5%. Although the steam injection will increase the total exergy losses, the exergy loss per MW output is much smaller than that of simple cycle.
- Presently, the STIG technology is applied only to the gas turbine cycle. It can be extended to the combined power-cycle which are now well established. The present study can be further extended for thermoeconomic or exergoeconomic analysis.

3. References

1 Nishida K, Takagi T, Kinoshita S, Regenerative steam-injection gas turbine systems, *Applied Energy* **81** (2005) 231-246; Japan.

2 Hawaj O M, Mutairi H, A combined power cycle with absorption air cooling, *Energy* **32** (2007) 971-982, Kuwait.

3 Pelster S, Favrat D,Von Spakovsky M R, The thermo economic analysis and environomic modeling and optimization of the synthesis and operation of combined cycle with advanced options, *Engineering for gas turbine and power, transaction of the ASME* **123** (2001) 717-26.

4 Bhargava R & Meher-Homji C B, Parametric analysis of existing gas turbines with inlet evaporative and overspray fogging, *Journal of Engineering for gas turbines and power* **127** (2005) 145; Houston.
5 Chaker M, Homji C B M, Mee I I I T, Inlet fogging of gas turbine engines-part II: fog droplet sizing analysis, nozzle types, measurement and testing, *Journal of engineering for gas turbines and power* **126** (2004) 559, Monrovia.
6 Bansode S, Sinha R, A thermodynamic analysis for gas turbine power optimization by fog cooling system, *20th national and 9th International ISHMT-ASME heat and mass transfer conference* (2010).
7 Alexis G K, Performance parameters for the design of a combined refrigeration and electrical power cogeneration system, *International journal of refrigeration* **30** (2007) 1097-1103, Greece.
8 Kumar A, Kachhwaha S S, Mishra R S, Thermodynamics analysis of a regenerative gas turbine cogeneration plant, *Journal of Scientific & Industrial Research*, **69** (2010) 225-231; India.
9 Wang F J & Chiou J S, Integration of steam injection and inlet air cooling for a gas turbine generation system, *Exergy conversion and Management*, **45** (2004) 15-26, Taiwan; ROC.
10 Bilgen E, Exergetic and engineering analysis of gas turbine based cogeneration systems, *Energy* **25** (2000) 1215-1229, Canada.
11 Ondryas I S, Wilson D A, Kawamoto M, Haub G L, Options in gas turbine power augmentation using inlet air chilling, *Journal of Engineering for gas turbines and power* **113** (1991) 205.

International Journal of Computer Applications (0975 - 8887)
Volume 58 - No. 5, November 2012

Yagi-Uda Antenna Design Optimization for Maximum Gain using different BBO Migration Variants

Satvir Singh
SBS State Technical Campus
Ferozepur, Punjab [INDIA]

Gagan Sachdeva
SBS State Technical Campus
Ferozepur, Punjab [INDIA]

ABSTRACT

Biogeography is the study of distribution of biological species, over space and time, among random habitats. Recently developed Biogeography-Based Optimization (BBO) is a technique, where solutions of the problem under consideration are named as *habitats*; similar to *chromosome* in Genetic Algorithms (GAs) and *particles* in Particle Swarm Optimization (PSO). Feature sharing among various habitats, i.e., *exploitation*, is made to occur due to migration operator wheras *exploration* of new SIV values, similar to that of GAs, is accomplished with mutation operator. In this paper, various migration variants of BBO algorithm, reported till date, are investigated to optimize the lengths and spacings for Yagi-Uda antenna elements for maximum gain. The results obtained with these migration variants are compared and the best results are presented in the ending sections of the paper.

Keywords:

Yagi-Uda Antenna, Bio-geography Based Optimization (BBO), Migration Variants, Enhanced BBO, Immigration Refusal BBO, Blended BBOifx

1. INTRODUCTION

Antenna is an electrical device which forms an interface between free-space radiations and transmitter or receiver. The choice of an antenna depends on various factors such as gain, impedance, bandwidth, frequency of operation, Side Lobe Level (SLL), etc. A Yagi-Uda antenna is a widely used antenna design at VHF and UHF due to its high forward gain capability (typically, greater than 10dB), low cost and ease of construction. It is a parasitic linear array of parallel dipoles, one of which is energized directly by transmission line while others act as a parasitic radiators whose currents are induced by mutual coupling. The characteristics of Yagi-Uda antenna are affected by its geometric parameters.

A Yagi-Uda antenna was invented in 1926 by H. Yagi and S. Uda at Tohoku University [28] in Japan, however, published in English in 1928 [32]. Since then, continuous efforts have been put in improving the antenna design for higher gain, requisite impedance, minimum SLL and wider bandwidth using different optimization techniques based on traditional mathematical approaches [21, 3, 7, 22, 6, 5, 8] and modern heuristics of Artificial Intelligence (AI) [13, 31, 30, 2, 15, 27, 26].

In 1949, Fishenden and Wiblin [12] proposed an approximate design of Yagi aerials for maximum gain. In 1959, Ehrenspeck and Poehler have given a manual approach to maximize the gain of the antenna by varying various lengths and spacings of its elements [11]. Later, with the availability of computer software at affordable prices made it possible to optimize antennas, numerically. Bojsen *et al.* [3] proposed another optimization technique to maximize the gain of Yagi-Uda antenna arrays with equal and unequal spacings between adjoining elements. Cheng *et al.* [6, 5] have used optimum spacings and lengths to optimize the gain of a Yagi-Uda antenna. In [8], Cheng has proposed optimum design of the antenna where antenna gain function is highly non-linear.

In 1975, John Holland introduced Genetic Algorithms (GAs) as a stochastic swarm based AI technique, inspired from Darwinian theory of *Survival of the fittest*, to optimize arbitrary systems for certain fitness measurement function. Many researchers, since then, investigated GAs to optimize Yagi-Uda antenna designs for various objectives one-by-one [1, 13, 9] and collectively [31, 29, 14, 24]. Baskar *et al.*, in [2], have optimized the antenna using Comprehensive Learning PSO (CLPSO) and presented better results than that of previously reported techniques. Li in [15] has used Differential Evolution (DE) to optimize geometrical parameters of a Yagi-Uda antenna and illustrated capabilities of the proposed method with several antenna designs. In [27], Singh *et al.* have explored another useful, stochastic global search and optimization technique named as Simulated Annealing (SA) for optimal Yagi-Uda antenna design.

In 2008, Dan Simon introduced yet another swarm based stochastic optimization technique, named BBO [23], inspired from the science of biogeography. Here, features sharing among various habitats, i.e., potential solutions, is accomplished with migration operator and exploration of new features is done with mutation operator. Singh *et al.* have presented BBO as a better alternative for optimizing Yagi-Uda antenna designs [26].

In [10], Du *et al.* have proposed the concept of immigration refusal in BBO aiming at improved performance. In [16], Ma and Simon introduced another migration operator, i.e., Blended migration, to solve constrained optimization problems and make BBO convergence faster. In [19], Pattnaik *et al.* have proposed Enhanced Biogeography-Based Optimization (EBBO) in which duplicate habitats, created due to migration of features, is replaced with randomly generated habitats to increase the exploitation ability of BBO algorithm.

In this paper, various migration variants in BBO algorithm are experimented to optimize the wire-lengths of Yagi-Uda antenna elements and their spacings to investigate relative optimization performance. A method of moments based programme, NEC2 (Numerical Electromagnetics Code version 2), is used to evaluate the antenna designs for gain, input impedance, bandwidth and beamwidth, etc.

After this brief literature review, the paper is outlined as follows: In the Section 2, Yagi-Uda antenna design parameters are discussed. Section 3 is dedicated to BBO algorithm and its variants. In Section 4, comparative simulation results obtained with different migration operators, during optimizing antenna designs, are presented. Finally, drawn conclusions and future scope have been discussed in Section 5.

2. THE ANTENNA DESIGN PARAMETERS

Yagi-Uda antenna consists of three types of elements: (a) *Reflector*–biggest among all and is responsible for blocking radiations

in one direction. (b) *Feeder*–which is fed with the signal from transmission line to be transmitted and (c) *Directors*–these are usually more then one in number and responsible of unidirectional radiations. Figure 1 depicts a typical six-element Yagi-Uda wire antenna with four directors parallel to x-axis and placed on y-axis. An incoming field sets up resonant currents on all the antenna elements which re-radiate signals. These re-radiated fields are then picked up by the feeder element, that leads to total current induced in the feeder equivalent to combination of the direct field input and re-radiated contributions from the director and reflector elements.

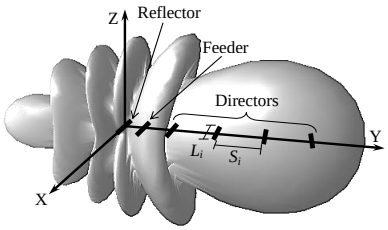


Fig. 1. Six-element Yagi-Uda Antenna

Designing a Yagi-Uda antenna involves determination of wire-lengths and wire-spacings in between to get maximum gain, desired impedance and minimum SLL at a frequency of operation. Typically, an antenna with N elements requires $M = 2N - 1$ parameters, i.e., N wire-lengths and $N - 1$ their spacings, that are to be determined. These M parameters, collectively, are represented as a string referred as a *habitat*, H, given as (1) or (2).

$$H = [L_1, L_2, \ldots, L_N, S_1, S_2, \ldots, S_{N-1}] \quad (1)$$

or

$$H = [L_1, S_1, \ldots, L_i, S_i, L_{i+1}, \ldots, S_{N-1}, L_N,] \quad (2)$$

where L_i is the length of ith element and S_i is the spacing between the ith and $(i + 1)$th element.

3. BIOGEOGRAPHY-BASED OPTIMIZATION

As name suggests, BBO is a population based global optimization technique developed on the basis of the science of biogeography, i.e., study of the distribution of animals and plants among different habitats over time and space. BBO results presented by researchers are better than other EAs [13, 29, 2, 20].

Mathematical models of biogeography describe migration, speciation, and extinction of species in various islands [17]. The term *island* is used for any habitat that is geographically isolated from other habitats. Habitats that are well suited residences for biological species are referred to have high Habitat Suitability Index (HSI) similar to fitness in other EAs. However, HSI value of a habitat depends upon many factors such as rainfall, diversity of vegetation, diversity of topographic features, land area, and temperature, etc. These variables that characterize habitability are termed as Suitability Index Variables (SIVs). The habitats with a high HSI tend to have a large population of its resident species, that is responsible for more probability of emigration and less probability of immigration due to natural random behavior of species. Immigration is the arrival of new species into a habitat or population, while emigration is the act of leaving one's native region.

On the other hand, habitats with low HSI tend to have low emigration rate, μ, due to sparse population, however, they will have high immigration rate, λ. Suitability of habitats with low HSI is likely to increase with influx of species from other habitats having high HSI. However, if HSI does not increase and remains low, species in that habitat go extinct that leads to additional immigration. For sake of simplicity, it is safe to assume a linear relationship between emigration rates and immigration rates. Additionally, same maximum emigration and immigration rates, i.e., $E = I$, as depicted graphically in Figure 2. For HSI of k-

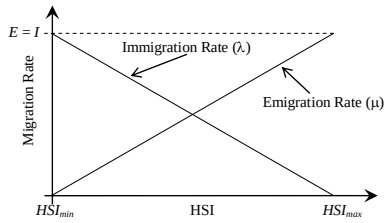


Fig. 2. Migration Curves

th habitat emigration and immigration rates are given by (3) and (4).

$$\mu_k = E \cdot \frac{HSI_k}{HSI_{max} - HSI_{min}} \quad (3)$$

$$\lambda_k = I \cdot \left(1 - \frac{HSI_k}{HSI_{max} - HSI_{min}}\right) \quad (4)$$

The immigration of new species from high HSI to low HSI habitats may raise the HSI of poor habitats as good solutions have more resistance to change than poor solutions whereas poor solutions are more dynamic and accept a lot of new features from good solutions. Following subsections discuss two constituents operators of BBO algorithm, viz. migration and mutation.

3.1 Migration

Migration is a probabilistic operator that improves HSI of poor habitats by sharing features from good habitats. During Migration, immigrating habitat, *ImHbt*, use its immigration rate, λ, given by (4), to probabilistically decide whether to immigrate or not. In case immigration is selected, then the emigrating habitat, *EmHbt*, is found probabilistically based on emigration rate, μ, given by (3). The process of migration is completed by copying values of SIVs from *EmHbt* to *ImHbt* at random chosen sites. This is termed as the standard migration operator, in this paper. The migration operator may lead to same types of habitats in large number after few iterations. In this paper, different migration variants are investigated to increase the diversity in the population, with objective of improved performance BBO algorithm whose pseudo code are depicted in Algorithm 3.1 and discussed in the following subsections. [b] Pseudo Code of Migration with

```
Select ImHbt with probability based on λ
if ImHbt is selected then
   Select EmHbt with probability based on μ
   if EmHbt is selected then
      Randomly select a SIV(s) from EmHbt
      switch (Option)
        case 1: Standard
           apply migration
        case 2: Immigration Refusal
           if (fitness(ImHbt)>fitness(EmHbt))
              apply migration
           end if
        case 3: Blended
           ImHbt(SIV) = α·ImHbt(SIV) + (1 – α)·EmHbt(SIV)
        case 4: EBBO
           apply migration
           eliminate duplicates
      end switch
   end if
end if
```

Variants

3.1.1 Immigration Refusal. In BBO, if a habitat has high emigration rate, i.e, the probability of emigrating to other habitats is high and the probability of immigration from other habitats is low. However, the low probability does not mean that immigration will never happen. Once in a while, a highly fit solution may receive solution features from a low-fit solution that may degrade its fitness. In such cases, immigration is refused to prevent degradation of HSI values of habitats. This BBO variants with conditional migration is termed as Immigration Refusal BBO [10] whose performance with testbed of benchmark functions is encouraging.

3.1.2 Blended Migration. Blended migration operator [16] is a generalization of the standard BBO migration operator that has got inspiration from blended crossover in GAs [18]. Here, SIVs of immigrating habitat, *ImHbt*, is not simply replaced by SIVs of emigrating habitat, *EmHbt*, like standard migration operation. Rather, a new value for i-th SIV, SIV_{new}, is comprised of two components given as (5)

$$SIV_{new} = \alpha \cdot ImHbt(SIV) + (1-\alpha) \cdot EmHbt(SIV) \quad (5)$$

where α is a random number between 0 and 1. The pseudo code of blended migration is depicted as case 3 in Algorithm 3.1.

3.1.3 Enhanced Biogeography Based Optimization. Standard BBO migration operator tends to create duplicate solutions which decreases the diversity in the population. To prevent this diversity decrease, duplicate habitats are replaced with randomly generated habitats, as depicted as case 4 in Algorithm 3.1. This leads to increase exploration of new SIV values. In EBBO, clear duplicate operator is integrated in basic BBO to improve its performance [19].

3.2 Mutation

Mutation is probabilistic operator that randomly modifies the values of some randomly selected SIVs. It is intended for exploration of search space for better SIV values by increasing the diversity in the population. Here, increased mutation rate is investigated on habitats those are, probabilistically, participating less in migration. The mutation rate, $mRate$, for i-th habitats is determined as (6)

$$mRate = C \cdot min(\mu_k, \lambda_k) \quad (6)$$

where C is a constant and equal to 3 [25]. The pseudo code for standard mutation operator is depicted in Algorithm 3.2. In our simulations, like other population-based optimization techniques, elitism is considered so as to preserve best solution throughout the process of optimization. [t] Generalized Pseudo Code for Mutation

$mRate = C \times \min(\mu_k, \lambda_k)$ where $C = 3$
for n = 1 to NP **do**
  for j = 1 to length(H) **do**
    Select H_j(SIV) with $mRate$
    If H_j(SIV) is selected **then**
      Replace H_j(SIV) with randomly generated SIV
    end if
  end for
end for

4. SIMULATION RESULTS AND DISCUSSIONS

Six-wire Yagi-Uda antenna designs are optimized for gain using BBO where all migration variants, discussed in Section 3.1, are investigated along with standard mutation operator. Average of 10 monte-carlo evolutionary runs for each migration operators are presented here for fair comparative study of stochastic algorithm variants, viz. immigration refusal, blended migration, EBBO and standard migration, as depicted in Figure 3. The C++ programming platform is used for coding of BBO algorithm, whereas, a NEC2 [4] is used for evaluation of antenna designs

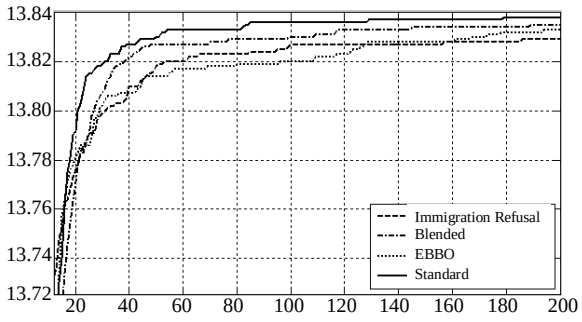


Fig. 3. Convergence Performance of BBO with Migration Options

Table 1. The best results obtained during gain optimization

	Standard		Immi. Refusal		Blended		EBBO	
Element	**Length**	**Spacing**	**Length**	**Spacing**	**Length**	**Spacing**	**Length**	**Spacing**
1(λ)	0.4832	-	0.4845	-	0.4838	-	0.4853	-
2(λ)	0.4794	0.2011	0.4927	0.1862	0.4689	0.1695	0.4888	0.1527
3(λ)	0.4410	0.2347	0.4398	0.2439	0.4414	0.2414	0.4411	0.2523
4(λ)	0.4273	0.3741	0.4234	0.3934	0.4265	0.3753	0.4268	0.3814
5(λ)	0.4205	0.4205	0.4201	0.4164	0.4181	0.4417	0.4185	0.4231
6(λ)	0.4294	0.3824	0.4256	0.3878	0.4274	0.3762	0.4236	0.3906
Gain(dBi)	**13.84**		**13.84**		**13.84**		**13.84**	
Z(Ω)	**4.78+j42.84**		**5.62+j64.60**		**3.58+j24.68**		**3.68+j54.11**	

based on method of moments. Each potential solution in BBO is encoded as vector with 11 SIVs as given by (0**??**). The universe of discourse for the search of optimum values of wire-lengths and wire-spacings are $0.40\lambda - 0.50\lambda$ and $0.10\lambda - 0.45\lambda$, respectively, however, cross sectional radius and segment sizes are kept same for all elements, i.e., 0.003397λ and 0.1λ respectively, where λ is the wavelength corresponding to frequency of operation, i.e, 300 MHz. Excitation is applied to the middle segment of driven element and location of middle segment of the reflector element is always kept at $x = 0$. Typically, the best antenna designs obtained during process of optimization are tabulated in Table 1.

EBBO algorithm yields poorest solutions in initial iterations and later improves significantly than immigration refusal variant. Blended migration variant gives mediocre performance. Comparatively, standard migration option gives the best convergence performance during optimizing Yagi-Uda antenna design for maximum gain in limited 200 iterations.

In our application of optimizing Yagi-Uda antenna design standard migration operator yields the best performance in limited 200 iterations. However, if run for large number of iterations EBBO may turn to be better option.

5. CONCLUSIONS AND FUTURE SCOPE

In this paper, BBO algorithm is applied to optimize six-element Yagi-Uda antenna designs for gain maximization. Here, BBO is experimented with different migration variants, viz., (1) Immigration Refusal, (2) Blended migration, (3) EBBO and (4) Standard migration, along with standard mutation operator. Here, it is observed that standard migration operator is the best option for optimizing the antenna designs. Investigation of BBO algorithms for other types of antenna designs is next on our agenda.

6. REFERENCES

[1] E.E. Altshuler and D.S. Linden. Wire-antenna Designs using Genetic Algorithms. *Antennas and Propagation Magazine, IEEE*, 39(2):33–43, 1997.

[2] S. Baskar, A. Alphones, P N Suganthan, and J J Liang. Design of Yagi-Uda Antennas using Comprehensive Learn-

ing Particle Swarm Optimisation. *IEEE*, 152(5):340–346, 2005.

[3] JH Bojsen, H. Schjaer-Jacobsen, E. Nilsson, and J. Bach Andersen. Maximum Gain of Yagi–Uda Arrays. *Electronics Letters*, 7(18):531–532, 1971.

[4] G. J. Burke and A. J. Poggio. Numerical Electromagnetics Code (NEC) method of moments. *NOSC Tech. DocLawrence Livermore National Laboratory, Livermore, Calif, USA*, 116:1–131, 1981.

[5] C. Chen and D. Cheng. Optimum Element Lengths for Yagi-Uda Arrays. *IEEE Transactions on Antennas and Propagation,*, 23(1):8–15, 1975.

[6] D. Cheng and C. Chen. Optimum Element Spacings for Yagi-Uda Arrays. *IEEE Transactions on Antennas and Propagation,*, 21(5):615–623, 1973.

[7] D. K. Cheng. Optimization Techniques for Antenna Arrays. *Proceedings of the IEEE*, 59(12):1664–1674, 1971.

[8] D. K. Cheng. Gain Optimization for Yagi-Uda Arrays. *Antennas and Propagation Magazine, IEEE*, 33(3):42–46, 1991.

[9] D. Correia, A. J. M. Soares, and M. A. B. Terada. Optimization of gain, impedance and bandwidth in Yagi-Uda Antennas using Genetic Algorithm. *IEEE*, 1:41–44, 1999.

[10] D. Du, D. Simon, and M. Ergezer. Biogeography-based Optimization Combined with Evolutionary Strategy and Immigration Refusal. *IEEE*, 1:997–1002, 2009.

[11] H. Ehrenspeck and H. Poehler. A New Method for Obtaining Maximum Gain from Yagi Antennas. *IRE Transactions on Antennas and Propagation,*, 7(4):379–386, 1959.

[12] R. M. Fishenden and E. R. Wiblin. Design of Yagi Aerials. *Proceedings of the IEE-Part III: Radio and Communication Engineering*, 96(39):5, 1949.

[13] E. A. Jones and W. T. Joines. Design of Yagi-Uda Antennas using Genetic Algorithms. *IEEE Transactions on Antennas and Propagation,*, 45(9):1386–1392, 1997.

[14] Y. Kuwahara. Multiobjective Optimization Design of Yagi-Uda Antenna. *IEEE Transactions on Antennas and Propagation,*, 53(6):1984–1992, 2005.

[15] J. Y. Li. Optimizing Design of Antenna using Differential Evolution. *IEEE*, 1:1–4, 2007.

[16] H. Ma and D. Simon. Blended Biogeography-based Optimization for Constrained Optimization. *Engineering Applications of Artificial Intelligence*, 24(3):517–525, 2011.

[17] R.H. MacArthur and E.O. Wilson. *The Theory of Island Biogeography*. Princeton Univ Pr, 1967.

[18] T. McTavish and D. Restrepo. Evolving Solutions: The Genetic Algorithm and Evolution Strategies for Finding Optimal Parameters. *Applications of Computational Intelligence in Biology*, 1:55–78, 2008.

[19] S. S. Pattnaik, M. R. Lohokare, and S. Devi. Enhanced Biogeography-Based Optimization using Modified Clear Duplicate Operator. *IEEE*, 1:715–720, 2010.

[20] M. Rattan, M. S. Patterh, and B. S. Sohi. Optimization of Yagi-Uda Antenna using Simulated Annealing. *Journal of Electromagnetic Waves and Applications, 22*, 2(3):291–299, 2008.

[21] D. G. Reid. The Gain of an Idealized Yagi Array. *Journal of the Institution of Electrical Engineers-Part IIIA: Radiolocation,*, 93(3):564–566, 1946.

[22] L. C. Shen. Directivity and Bandwidth of Single-band and Double-band Yagi Arrays. *IEEE Transactions on Antennas and Propagation,*, 20(6):778–780, 1972.

[23] D. Simon. Biogeography-based Optimization. *IEEE Transactions on Evolutionary Computation,*, 12(6):702–713, 2008.

[24] Satvir Singh, Eitika Mittal, and Gagan Sachdeva. Multi-Objective Gain-Impedance Optimization of Yagi-Uda Antenna using NSBBO and NSPSO. *International Journal of Computer Applications*, 56(15):1–6, 2012.

[25] Satvir Singh and Gagan Sachdeva. Mutation Effects on BBO Evolution in Optimizing Yagi-Uda Antenna Design. In *proceeding of IEEE*, 2012.

[26] U. Singh, H. Kumar, and T. S. Kamal. Design of Yagi-Uda Antenna Using Biogeography Based Optimization. *IEEE Transactions on Antennas and Propagation,*, 58(10):3375–3379, 2010.

[27] U. Singh, M. Rattan, N. Singh, and M. S. Patterh. Design of a Yagi-Uda Antenna by Simulated Annealing for Gain, Impedance and FBR. *IEEE*, 1:974–979, 2007.

[28] Shintaro Uda and Yasuto Mushiake. *Yagi-Uda Antenna*. Maruzen Company, Ltd, 1954.

[29] N. V. Venkatarayalu and T. Ray. Single and Multi-Objective Design of Yagi-Uda Antennas using Computational Intelligence. *IEEE*, 2:1237–1242, 2003.

[30] N.V. Venkatarayalu and T. Ray. Optimum Design of Yagi-Uda Antennas Using Computational Intelligence. *IEEE Transactions on Antennas and Propagation,*, 52(7):1811–1818, 2004.

[31] H. J. Wang, K. F. Man, C. H. Chan, and K. M. Luk. Optimization of Yagi array by Hierarchical Genetic Algorithms. *IEEE*, 1:91–94, 2003.

[32] H. Yagi. Beam Transmission of Ultra Short Waves. *Proceedings of the Institute of Radio Engineers*, 16(6):715–740, 1928.

7. AUTHOR'S PROFILE

Satvir Singh was born on Dec 7, 1975. He received his Bachelor's degree (B.Tech.) from Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab (India) with specialization in Electronics & Communication Engineering in year 1998, Masters degree (M.E.) from Delhi Technological University (Formerly, Delhi College of Engineering), Delhi (India) with distinction in Electronics & Communication Engineering in year 2000 and Doctoral degree (Ph.D.) from Maharshi Dayanand University, Rohtak, Haryana (India) in year 2011.

During his 12 years of teaching experience he served as Assistant Professor and Head, Department of Electronics & Communication Engineering at BRCM College of Engineering & Technology, Bahal, (Bhiwani) Haryana, India and as Associate Professor & Head, Department of Electronics & Communication Engineering at Shaheed Bhagat Singh State Technical Campus (Formerly, SBS College of Engineering & Technology), Ferozepur Punjab, India.

His fields of special interest include Evolutionary Algorithms, High Performance Computing, Type-1 & Type-2 Fuzzy Logic Systems, Wireless Sensor Networks and Artificial Neural Networks for solving engineering problems. He is active member of an editorial board of International Journal of Electronics Engineering and published nearly 15 research papers in International Journals and Conferences. He has delivered nearly 15 Invited Talks during National and International Conferences, Seminar, Short Term Courses and Workshops. He completed two AICTE funded projects under MODROB Scheme worth 15 Lacs and conducted a 2-Week Staff Development Programme on "Intelligent Computational Techniques".

Gagan Sachdeva was born on Oct 4, 1988. He received his Bachelor's degree (B.Tech.) and Master's degree (M.Tech.) from

Shaheed Bhagat Singh State Technical Campus (formerly, SBS College of Engineering & Technology), Ferozepur, Punjab (India) with specialization in Electronics & Communication Engineering in year 2010 and 2012, respectively. He is presently with Shaheed Bhagat Singh State Technical Campus, Ferozepur as Adhoc Lecturer in the department of ECE. His research interests include Evolutionary Algorithms, Antenna design optimization, and Wireless Sensor Networks.